THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY

CARL MURCHISON

VOLUME 57

1940

Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

TABLE OF CONTENTS

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price will be $7 per volume, two volumes per year.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any zinc or half tone cuts.

(6). The publication charges will be as follows:

4	page signature	$12
8	" "	24
12	" "	36
16	" "	48
20	" "	60
24	" "	72
28	" "	84
32	" "	96
36	" "	108
40	" "	120

These charges are tentative and will be adjusted as justified by the subsequent experience, especially by the effect of the war in China and in Europe.

(7). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, in accordance with publication needs.

Two volumes annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

Two volumes per year
QUARTERLY

September, 1940
Volume 57, First Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

SEPTEMBER, 1940

(Manuscripts are printed in the order of final acceptance)

(OVER)

Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass., under the Act of March 3, 1879

Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

TABLE 1

THE DISTRIBUTION OF INTELLIGENCE TEST SCORES IN A GROUP OF EIGHT HUNDRED PUPILS IN GRADES THREE TO SIX, INCLUSIVE, WHO ANSWERED THE CHECK LIST—BY SEX AND GRADE

	GRADES										
	III		IV		V		VI		Total		
IQ	*B*	*G*	*B*	*G*	*B*	*G*	*B*	*G*	*B*	*G*	Total
169-160						1	1		1	1	2
159 150	1			1	1	1		1	2	3	5
149-140		1	1	1		1	1	1	2	4	6
139-130	1	3	3	2	3	1	2	2	9	8	17
129-120	6	4	4	4	4	4	4	4	18	16	34
119-110	16	16	16	16	16	16	16	16	64	64	128
109-100	30	30	30	30	30	30	30	30	120	120	240
99-90	22	22	22	22	22	22	22	22	88	88	176
89-80	17	17	17	17	17	17	17	16	68	67	135
79-70	7	6	6	3	6	5	6	7	25	21	46
69-60		1	1	4	1	2	1	1	3	8	11
N =	100	100	100	100	100	100	100	100	100	100	800

B = Boy, *G* = Girl.

The range of chronological ages in each grade in terms of years and months:

Grade III,	7 1 to 10 2
Grade IV,	8.1 to 12 3
Grade V,	9 0 to 13 2
Grade VI,	10 2 to 13 5

few new personnel. The average age of the teachers was slightly more than 40 years All were regularly appointed teachers, had passed a uniform examination, and had been required to take "alertness" courses to attain their maximum salaries Their age warrants the inference that the majority had reached their maximum salary About one-half of them were in schools where the "activity program" was in operation, and were using to a greater or less extent the progressive methods advocated for this type of instruction The other teachers were using the approved methods of the conventional schools Both groups followed the required courses of study for the elementary grades of the New York City Public Schools. Student teachers under the direct supervision of the investigator were practicing in the class rooms where the investigation took place. These assistants were members of the senior class of a four year college, and had had previous experience in pupil testing.

Two instruments were used in the study The first was a preliminary questionnaire tried out on 800 children in Grades 3-6,

inclusive The second was a check list sheet called, for the purpose of interesting the children who were to use it, *A Funny Test* It is reproduced below as Form 1 This test was given to 800 children

FORM 1

A FUNNY TEST

Directions

Draw a line under *three* answers for each question If you do not think any are funny, write one that is Turn your paper over when you finish each question Wait for the teacher to read the questions Do not talk

I *The funniest people in the movies are*

1 Betty Boop
2 Charlie Chaplin
3 Donald Duck
4 Eddie Cantor
5. Joe E Brown
6 Joe Penner
7 Laurel and Hardy
8 Marx Brothers
9 Martha Raye
10 Mickey Mouse
11 Mae West
12 Popeye
13 Ritz Brothers

II *The funniest people on the radio are*

1 Burns and Allen (Gracie)
2. Charlie McCarthy
3 Eddie Cantor
4 Fibber McGee
5. Fred Allen
6 Jack Benny
7 Joe Penner
8 Parkyakarkus
9. Pick and Pat
10 Popeye
11 The Goldbergs

III *The funniest comic strips are*

1 Bringing Up Father
2 Blondie
3 Donald Duck
4 Henry
5 Katzenjammer Kids
6 Moon Mullins
7 Mickey Mouse
8 Popeye
9 Smitty
10 Smokey Stover
11 The Goldbergs
12 Winnie Winkle

IV *Children make me laugh when they*

1 Act in school plays
2 Are stupid
3 Fight
4 Get punished
5 Laugh, so I laugh too
6 Make faces
7 Make mistakes
8 Make noises
9 Stutter
10 Talk out in school
11 Tell jokes or riddles
12 Tickle me

V *Grown people make me laugh when they*

1 Ask silly questions
2 Are drunk
3 Dance
4 Fall down
5 Get very angry
6 Kiss each other
7 Make mistakes
8 Make faces
9 Play tricks
10 Sing funny songs
11 Smoke
12 Tell jokes or stories
13 Tickle me

VI. *People look funny when they*

1. Are very fat
2. Are very thin
3. Have big ears
4. Have big noses
5. Have no arms
6. Have cross eyes
7. Shoot each other
8. Wear funny hats
9. Wear spectacles
10. Are in accidents
11. Fall down
12. Sing

VII *Things that make me laugh are*

1. Cats
2. Chinese people
3. Cowboys
4. Clowns
5. Dunces
6. Jack-o'-lanterns
7. Queer words
8. Girls in boys' clothes
9. Boys in girls' clothes
10. Children jumping rope
11. Acrobats
12. Houses with queer windows

VIII *Answer one*

A funny story is . . . . .

IX *I read it*

1. In my reader
2. In my own book
3. In a library book

X *It is funny because*

1. It uses queer words or sounds
2. It is impossible
3. It surprises you
4. Animals act like people
5. It has funny pictures
6. It is foolish
7. People would not act like those in the story

XI *A funny poem is* . . . .

1. I heard it in Grade . . . .
2. I heard it outside of school

XII *A funny song is* . . . .

1. I heard it at school
2. I heard it on the radio
3. I heard it at a movie
4. I heard it at . . . .

XIII *If you tell jokes, where do you get them?*

1. Joke books
2. Radio
3. Other children
4. People at home
5. Magazines
6. Newspapers

XIV *Do you laugh*

1. Often
2. Sometimes
3. Almost never

XV *Answer as many as you wish*

Does your teacher make you laugh when she

1. Tells funny stories
2. Reads funny stories
3. Punishes other children
4. Forgets
5. Makes mistakes
6. Laughs often herself
7. Says sarcastic things
8. Tells jokes
9. Gets angry
10. Makes queer motions

XVI *Did you think this test was fun?*

1. Yes
2. No

XVII *Turn your paper over Draw a funny picture Tell why it is funny*

of other Grades 3-6, inclusive The *Funny Test* was administered by student teachers who had been trained to give it. Each item was read aloud by the one giving the test, while the pupils marked their papers. This avoided the factor of reading variability, while the checking procedure eliminated the varying abilities of the children in writing and spelling The students who gave the test reported that almost without exception the pupils thoroughly enjoyed it For example:

> "I gave two 'Funny Tests' (5-*A*, 5-*B*) The test made a 'big hit' with the children, who want to know if I have 'any more of those nice tests'"
>
> "Most of them liked the lesson very much, and it *is* the first time I have seen them so animated You know it is a very slow class The teacher had some difficulty bringing them back to their usual order after we had finished."
>
> "Children laughed a great deal, and were almost hysterical at times Some of them were perplexed when they couldn't think of some of the answers, but on the whole a spirit of fun dominated during the period"

Question may be raised as to the small number of choices provided in the check list and the explanation is an important one The choices provided were those that the preliminary questionnaire had shown to be of most importance to the children themselves It would obviously have been impracticable to have had these young children consider long lists of items. The questionnaire indicated that the principal items were few and would probably be covered by the ones selected for the check list.

This report gives a brief summary and discussion of the data relating to the movies, radio, comics, and other items marked on the check list, and some of the data from the questionnaire.

For purposes of comparison, a group of 100 men and women students in graduate courses in a department of education of a large city university answered the items on the check list

The tables presenting the data shown hereinafter are by grades and by sexes, and are in terms of percentages except where otherwise indicated The bases for computing the percentages were the total number of replies for the various items. These numbers varied considerably because, while the children were asked to make a certain number of replies to each item, some of them made more and some fewer replies than directed In tabulating the results, all the replies made were used

B. Results

1. *Movies*

Table 2 gives the names of all the movie and radio persons reported as funny by the 800 children who answered the preliminary ques-

TABLE 2

Motion Picture and Radio Stars and Programs Considered Funny by Eight Hundred Pupils Who Answered the Questionnaire

(Number of times mentioned is given in parenthesis.)

Armetta, Henry (1)	Lum and Abner (2)
Allen, Fred (13)	Livingston, Mary (2)
Allen, Gracie (65)	Laurel and Hardy (41)
Allen, Jimmy (1)	Loy, Myrna (1)
Amos and Andy (2)	Lillie, Beatrice (1)
Armstrong, Jack (1)	Lloyd, Harold (3)
Astaire, Fred (1)	Little Orphan Annie (6)
Auer, Mischa (1)	Marshall, Herbert (1)
Ameche, Don (2)	McGee, Fibber (9)
Baker, Phil (4)	Marx Brothers (99)
Benny, Jack (124)	Moe, Larry and Curley
Beery, Wallace (1)	Moore, Victor (1)
Bergen, Edgar (6)	Mickey Mouse (62)
Bernie, Ben (1)	Murray, Ken (2)
Berle, Milton (1)	Morgan, Helen (1)
Bing, Herman (1)	Morgan, Frank (2)
Blondell, Joan (1)	News Report (1)
Boland, Mary (2)	O'Keefe, Walter (1)
Boop, Betty (24)	Oliver, Edna (1)
Bowes, Major (9)	Okey, Jack (1)
Brown, Joe E (46)	Omar the Mystic (1)

TABLE 2 (*continued*)

MOTION PICTURE AND RADIO STARS AND PROGRAMS CONSIDERED FUNNY BY EIGHT HUNDRED PUPILS WHO ANSWERED THE QUESTIONNAIRE
(Number of times mentioned is given in parenthesis.)

Brice, Fannie (3)	Pick and Pat (21)
Burns, Bob (2)	Parkyakarkus (35)
Burns, George (65)	Popeye (56)
Butterworth, Charles (2)	Penner, Joe (92)
Cagney, James (1)	Pitts, Zazu (4)
Cook, Joe (2)	Summerville, Slim (1)
Chaplin, Charlie (89)	Pierce, Al (1)
Cantor, Eddie (147)	Ritz Brothers (123)
Crosby, Bing (1)	Lone Ranger (2)
Chase and Sanborn Hour (19)	Rogers, Buck (1)
Charlie McCarthy (36)	Robinson, Bill (2)
Chan, Charlie (3)	Ripley, Robert (1)
Donald Duck (33)	Ruggles, Charlie (3)
Durante, Jimmy (4)	Rooney, Mickey (2)
Davis, Joan (1)	Raye, Martha (62)
Easy Aces (2)	Ross, Lanny (1)
Eno Crime Clues (5)	Reading the Funnies (3)
Farrell, Glenda (2)	Step'n Fetchit (1)
Fat and Skinny (25)	Stoopnagle and Bud (1)
Felix the Cat (2)	Show Boat (1)
Fields, W. C. (6)	Stroud Twins (2)
Faye, Alice (1)	Sally Lou (1)
Garbo, Greta (1)	Sparks, Ned (1)
Gable, Clark (2)	Tim and Irene (1)
Goldbergs (2)	Three Stooges (41)
Gang Busters (3)	Temple, Shirley (9)
Gibson, Hoot (1)	Tracy, Dick (7)
Horton, Edward E. (7)	Taylor, Robert (1)
Herbert, Hugh (12)	Tarzan (2)
Hit Parade (1)	Uncle Don (1)
Haley, Jack (1)	Vallee, Rudy (1)
Healy, T. (1)	Wynn, Ed (3)
Jolson, Al (13)	Withers, Jane (11)
Kelly, Kitty (1)	Wheeler and Woolsey (3)
Kelly, Patsy (8)	West, Mae (16)
Krazy Kat (1)	Wallington, James (2)
Kibbee, Guy (1)	Witch's Tale (1)

tionnaire. The number of children who named each person is also given. One hundred fifteen characters were named. Twenty-two of them were named by more than 12 children each. These are listed below, together with the numbers of children choosing each. *M* stands for movie star, and *R* for radio star. Fibber McGee and the Goldbergs were added as a check to see if the choices of the second group of 800 children agreed roughly with the choices of the first

800. It turned out that the results of the two groups were very similar.

M. R	Eddie Cantor	147	M	Joe E. Brown	46
R R.	Jack Benny	124	M.	Three Stooges	41
M	Ritz Brothers	123	M	Laurel & Hardy	41
M.	Marx Brothers	99	R.	Parkyakarkas	35
M R	Joe Penner	92	M.	Donald Duck	33
M	Charlie Chaplin	89	M	Fat & Skinny	25
R	Burns & Allen	65	M	Betty Boop	21
M R	Mickey Mouse	62	R	Pick & Pat	21
M R	Martha Raye	62	M	Mae West	16
R M	(Charlie McCarthy 36,		R	Fred Allen	13
	(Edgar Bergen 6,		R	Al Jolson	13
	(Chase & Sanborn 19,	61	R	Fibber McGee	9
M. R.	Popeye	66	R	Goldbergs	2

Several of these were both radio and movie stars. In the event that this was true, their names were included on both the movie and radio lists which the second group of pupils marked on the check list.

TABLE 3

PERCENTAGES OF CHOICES OF FUNNIEST MOVIE STARS BY SEX AND GRADES

	Grades								Teachers
	Boys				Girls				
	3	4	5	6	3	4	5	6	
a—Both sexes increasing									
Ritz Brothers	10	14	14	19	10	17	18	18	11
Laurel & Hardy	7	13	17	22	6	9	17	18	10
b—Both sexes decreasing									
Joe E Brown	12	5	9	4	9	6	8	5	10
Mickey Mouse	10	5	3	3	10	6	6	2	8
Betty Boop	6	2	2	0	10	4	3	2	1
c—Boys increasing—girls little change									
Marx Brothers	6	10	18	16	6	7	7	7	11
d—Boys decreasing—girls increasing									
Donald Duck	14	10	6	5	9	8	11	16	7
e—Boys little change—girls decreasing									
Charlie Chaplin	13	12	13	11	12	11	8	6	13
f—Little change both sexes									
Eddie Cantor	8	11	8	10	9	12	6	10	8
Joe Penner	4	9	4	4	7	8	6	6	3
Popeye	4	5	2	2	5	3	7	3	7
Martha Raye	4	2	2	2	5	9	4	6	3
Mae West	1	1	0	1	3	1	0	1	1
Others									5

Charlie McCarthy was not put in the movie group because at the time the study was made his motion picture career had not yet begun. Al Jolson's name was omitted from the check list because it was decided that he was more strictly a singer than a movie or radio humorist, although 13 children named him as a funny star.

Table 3 shows by grades and sexes the selections of the movie characters thought to be the funniest ones the children knew. The most popular characters, generally speaking, were the Ritz Brothers, Laurel and Hardy, Charlie Chaplin, and the Marx Brothers The popularity of choices seemed to be quite varied, however, both by sexes and grades. The arrangement of the names in the table brings out some of these variations. Probably some of the apparent trends seen are not true trends, because of the inevitable unreliabilities in the data, but interesting observations may be made. The increases in popularity of the Ritz Brothers and of Laurel and Hardy by both older boys and girls seems to be unmistakable. So do the decreases in choices of Mickey Mouse, Betty Boop, and possibly Joe E Brown. One of the most interesting differences is in regard to Donald Duck The older boys distinctly lost interest in him, while the older girls seemed to have increasing interest. Charlie Chaplin held a substantial place of affection in the hearts of all groups of the boys, but his glamour apparently faded among older girls The Marx Brothers likewise showed interesting trends. For the girls of all groups, they were secondary humor, but among the boys, there seemed to be a rapid rise of preference with age, to place them among the top ranking humorists

For five characters there seemed little variation from grade to grade or between the sexes Of these, Eddie Cantor held a high place for all groups. Joe Penner, Popeye, and Martha Raye were less popular, and Mae West was very low.

Comparison of the children's choices with those of the 100 teachers shows some interesting points. First, there seemed to have been some points of general agreement For example, both children and teachers put the Ritz Brothers, Laurel and Hardy, Charlie Chaplin, the Marx Brothers, and Eddie Cantor at the top of the groups of funny characters The percentages of older children's choices, however, for the Ritz Brothers and for Laurel and Hardy were much higher than those of the teachers It looks as if 3rd and 4th grade children agreed with adult teachers, while 6th grade pupils saw

much more humor in these characters than did the teachers. The same trend was indicated by the boys regarding the Marx Brothers and by the girls regarding Donald Duck. On the other hand, the adults showed a fondness for Mickey Mouse, Popeye, and Joe E. Brown that the 6th grade children seemed to have far "outgrown"! Apparently, the appreciation of humor is a complicated matter. Probably, through analysis it would show that the adults thought Mickey Mouse and Popeye were funny for quite different reasons than did the children. Very possibly too, the 6th graders were amused by some qualities of the Ritz Brothers that did not seem so funny to many teachers. It is probable, also, that the 3rd graders were intrigued by still other qualities than the 6th graders or the adults. That boys and girls are differently affected is self evident in the instances showing such varying trends, namely Donald Duck, Charlie Chaplin, and the Marx Brothers. In fact, this line of reasoning would mean that basically individual differences in humor are the rule. These conclusions seem to be supported by study of later tables.

2. *Radio*

Table 4 gives the percentages of choices by sex and grade for the three funniest radio stars. The most popular characters in general were, first and foremost for all groups of children, Charlie McCarthy, second, Parkyakarkus; third, Eddie Cantor; fourth, Burns and Allen.

The trends by grades and sexes showed some interesting variations. An increase with age seemed to be found for Burns and Allen and a decrease with age for Popeye. Boys showed decreasing choices with age for Jack Benny and Joe Penner, while the girls' interests remained about the same. The older boys also showed some increase for Eddie Cantor, while his popularity with older girls remained about the same as with younger.

The remaining six characters showed little or no change with the ages of the children choosing them. Among these were the two leading favorites, Charlie McCarthy and Parkyakarkus. The other four radio characters had very low percentages and would seem to be much less important. They were Fibber McGee, Pick and Pat, Fred Allen, and the Goldbergs.

Comparison of the teachers' choices with the children's preferences

TABLE 4

PERCENTAGES OF CHOICES OF FUNNIEST RADIO STARS BY SEX AND GRADES

	Grades								
	Boys				Girls				Teachers
	3	4	5	6	3	4	5	6	
	a—Both sexes increasing								
Burns & Allen	9	12	14	12	11	6	12	19	20
	b—Both sexes decreasing								
Popeye	7	4	3	1	8	[illegible]	3	1	2
	c—Boys decrease, girls little change								
Jack Benny	15	10	11	7	11	10	8	9	19
Joe Penner	15	12	10	9	9	8	11	10	3
	d—Boys increase, girls little change or decrease								
Eddie Cantor	12	14	12	24	12	9	12	8	7
	e—Little or no change both sexes								
Charlie McCarthy	18	20	20	21	23	21	25	22	22
Parkyakarkas	13	19	14	16	17	19	17	19	8
Fibber McGee	5	2	6	5	3	5	5	4	4
Pick & Pat	5	5	4	2	3	3	4	4	—
Fred Allen	2	1	4	3	3	2	4	2	11
Goldbergs	0	0	1	0	3	1	1	1	2
Others									2

brings out similarities and contrasts First, Charlie McCarthy was at the top for teachers as well as children and by practically the same percentage The teachers' next two favorites, however, Burns and Allen and Jack Benny, were not similarly esteemed by all the children's groups The 6th grade girls gave practically the same percentage for Burns and Allen as the teachers did, but not the boys For Jack Benny both boys and girls showed considerably less interest, the boys a decreasing interest with age The fourth highest choice of the teachers, Fred Allen, was one of the lowest of the children's The percentage of teacher choice for Parkyakarkus was about half that of the children, that for Joe Penner from a third to a fifth of the children's and among the lowest of the teachers' percentages Eddie Cantor, who stood among the top four with the children, was well down in the choices of the teachers Fibber McGee and the Goldbergs had relatively little humor for the teachers as well as for the children. Finally, the teachers showed practically the same interest as the older children in Popeye, the percentage choices being two per cent and one per cent

Doubtless, a better knowledge of the qualities in the radio persons which children and adults think funny would show that in this group of characters different things were held to be funny by different groups of boys, girls, and adults, as was pointed out above in connection with the movie characters. For example, the percentage of boys choosing Eddie Cantor increased from 12 per cent in Grade 3 to 24 per cent in Grade 6, as compared with 19 per cent of teachers who selected him. Both of these trends by the boys were away from the standard of the adults, but in opposite directions. This can hardly be explained on other grounds than that different qualities were thought funny by the various age and sex groups.

This probable fact would help to explain also the differences in interest in two stars who appeared in both the movie and radio lists, Eddie Cantor and Joe Penner. For the boys Eddie Cantor had a fairly high and unchanging percentage in all four grades. As a radio star, however, there was a great increase in choice by the older boys. Joe Penner's position as a movie star was low and changed little, if any, among the boys of the four grades. As a radio character, however, he was very popular with the 3rd grade boys but lost some of that preference in the 4th, 5th, and 6th grades. On the other hand, all the girls and also the teachers groups indicated about the same interest in Cantor and Penner as movie or radio characters.

3. *Comics*

Table 5 shows the percentages for favorite comics. The Katzenjammer Kids clearly lead the list and outclass all others in gain and popularity with age. The leading position of this old and and venerable comic is almost a unique case these days in which tradition rarely wins a kind word.

Next in popularity, and also increasingly so with age, is Henry. Donald Duck was in the top group for 3rd grade children, but his popularity waned considerably with the older pupils. Mickey Mouse stood up well except with older girls. Smokey, Smithy, Popeye, Moon Mullins, Blondie, Winnie Winkle, and Bringing Up Father bring up the list of choices with fairly large percentages each. The Goldberg check received but scattering preference.

Three sex differences seem to be real ones. Winnie Winkle, Blondie, and Bringing Up Father were more popular with girls in

TABLE 5

PERCENTAGES OF CHOICES OF FUNNIEST COMICS BY SEX AND GRADES

	Grades								
	Boys				Girls				Teachers
	3	4	5	6	3	4	5	6	
	a—Both sexes increasing								
Henry	13	16	14	19	12	13	13	15	5
Katzenjammer Twins	17	19	24	22	13	14	22	25	9
	b—Both sexes decreasing								
Popeye	8	6	4	3	8	5	7	2	9
Donald Duck	21	15	11	9	13	12	9	5	9
Moon Mullins	7	3	2	2	8	6	3	4	9
	c—Boys increase, girls no change								
Smokey	6	10	12	12	7	6	8	6	2
	d—Girls increase, boys no change								
Blondie	5	5	7	6	5	9	10	13	9
	e—Girls decrease; boys no change								
Mickey Mouse	11	10	8	10	11	8	5	2	9
	f—Little or no change, both sexes								
Smitty	6	5	7	8	5	8	7	6	6
Winnie Winkle	5	6	4	5	10	10	9	10	5
Bringing Up Father	1	4	7	4	8	7	8	11	22
Goldbergs	0	1	0	0	0	1	0	1	0
									94
Others									6

every group, except Blondie for whom, in Grade 3, both boys and girls gave five per cent choice In all of these comics there are dominant feminine characters. The balancing interest of the boys was divided between Donald Duck, Smokey, the Katzenjammer Kids, and Mickey Mouse

Comparison with the teachers' choices shows greater differences than in the data for either the movies or the radio. The percentage of teachers choosing the Katzenjammer Kids was only nine compared with 24 per cent and 25 per cent by 5th grade boys and 6th grade girls. On the other hand, the first choice of the teachers—Bringing Up Father—was by a percentage of 22 compared with percentages of from 1-11 for the children's groups The comic least popular in the opinion of the teachers was Smokey which showed but two per cent of choices, compared with 6-12 per cent by the children. Don-

ald Duck, one of the highest favorites of the children, was also one of the second highest group of adult favorites, along with Mickey Mouse, Popeye, the Katzenjammer Kids, Moon Mullins, and Blondie. Winnie Winkle received five per cent of the adults' choices, practically what the boys gave her, but about half of the girls' choices

A final interesting comparison may be seen between adults and children. Both the boys and girls showed decreasing interest in Popeye and Moon Mullins and the girls in Mickey Mouse The percentage fell to such a degree that by the 6th grade the children had percentages for these comics much below the percentages given by the adults It would be interesting to know where and when a later rise in children's interests would occur to bring the percentages up again, or if the selection of the teacher group were the chief factor causing the differences.

A further comparison may be made between choices of certain characters which were in more than one of these groups of characters Donald Duck is in both the movie and the comics In the former it was found that his popularity decreased with age for the boys from quite high to low and increased from a fairly high to very high place for the girls In the comics, on the other hand, the percentages for both boys and girls fell sharply from the highest percentages of all in Grade 3 to very low in Grade 6. Why did the same girls by grades give Donald Duck larger and larger percentages as a movie humorist, but show a declining choice for his humor in the comics? Why did the boys show decreasing interest by grade for both comic and movie Donald Duck? These questions do not seem answerable on the basis of the data.

Popeye's name was found in all the groups—movie, radio, and comics The trends of interest were decreasing in all his manifestations for both sexes, except for his movie rôle, which was quite low for all groups In the radio and comics the percentages fell from about seven or eight to about two in Grade 6

4. *Children Make Me Laugh*

Table 6 shows the percentages for the 12 items on the check list for what children do that amuses other children The four funniest things other children do were, in order, telling jokes, tickling, making faces, and laughing when others laughed. The first of these

TABLE 6

PERCENTAGES OF CHOICES OF FUNNIEST THINGS CHILDREN DO BY SEX AND GRADES

	Grades								
	Boys				Girls				Teachers
	3	4	5	6	3	4	5	6	
	a—Both sexes increasing								
Tell jokes	17	20	22	29	17	23	18	27–	28
	b—Both sexes decreasing								
Tickle me	30	26	26	17	30	27	29	24	7
	c—Little or no change, both sexes								
Make faces	16	16	14	18	18	17	15	20	15
Laugh	12	9	10	10	8	8	14	10	26
Make noises	6	5	3	6	4	4	4	3	2
Act in schl play	5	5	4	5	4	6	4	5	10
Are stupid	4	5	7	2	6	6	5	2	0
Stutter	4	3	6	5	3	4	4	4	0
Talk out in sch	1	3	2	2	2	1	3	3	2
Fight	2	2	1	1	1	1	0	0	1
Get punished	1	2	2	1	4	1	1	0	1
Make mistakes	0	3	3	4	2	1	1	1	5
Other									3

shows an unmistakable increase for both boys and girls as they grow older. The second shows a distinct decrease for boys, not so definitely a decrease for girls Making faces and laughing seemed to change little with age All the other eight causes for laughter had quite small or very small percentages for all groups regardless of age or sex

Comparison with the teachers' responses is worthwhile only when keeping in mind the greatly differing meanings of the question when put to children and when put to teachers With that understood it may be noted that teachers said they were chiefly amused by children when the latter told jokes or laughed These two causes account for over half of the declared causes for the teachers laughing *with* (it would seem) and not *at* the children This is a healthful state of affairs, if true There may be added to this 56 per cent, another 15 per cent due to children's making faces (wry, perplexed faces or ugly faces at teacher or others is not revealed) and 10 per cent for acting in school plays This makes a total of 81 per cent of teacher laughing probably *with* children rather than *at* them May this or even better be true!

5. *Grown People*

Table 7 shows the results regarding laughter produced by what grown people do The three main incentives to laughter seemed to

TABLE 7
PERCENTAGES OF CHOICES OF FUNNIEST THINGS GROWN PEOPLE DO BY SEX AND GRADES

	Grades								Teachers
	Boys				Girls				
	3	4	5	6	3	4	5	6	
	a—Both sexes increasing								
Play tricks	5	8	10	11	6	10	8	11	7
Tell jokes	6	9	13	16	7	13	14	14	26
Sing funny songs	7	10	12	11	5	11	13	14	24
	b—Both sexes decreasing								
Kiss each other	24	18	14	13	22	13	13	8	1
Tickle me	21	18	17	13	22	17	18	17	5
	c—Boys decreasing, girls little change								
Fall down	10	4	4	3	6	7	5	7	3
	d—Both sexes little or no change								
Drunk	12	16	14	12	12	14	15	12	8
Ask silly questions	5	6	5	8	4	7	6	7	11
Make faces	6	6	6	5	8	5	4	5	4
Dance	2	1	1	2	4	1	2	2	1
Very angry	1	1	1	2	1	0	1	2	7
Make mistakes	1	1	2	4	1	1	1	1	3
Smoke	1	2	0	1	2	0	0	0	0
Others									2

have been grown people tickling children, adults kissing each other, and drunken people. The last two seem to be affairs strictly between adults, quite unrelated to the fact that children might be concerned. Yet next to being tickled these two forms of behavior struck the children in Grades 3 to 5 as being the funniest things adults do! The next three things that grown people do which these children thought funny were telling jokes, singing funny songs, and playing tricks. Older children thought these things funnier than younger. The remaining items gave small percentages

Very insignificant sex differences are shown in the table. Older boys seemed to think that grown people falling down was less amusing that did younger boys. That may have been slightly more mirth provoking for girls than for boys of Grades 4-6. Grown people

kissing each other seemed less funny to sixth grade girls than to boys

The teachers' percentages show striking differences from the children's. Jokes and funny songs give one-half the causes for teachers laughing The children's answers showed an increasing importance of these two causes with increasing age The things that younger children felt funniest—people kissing each other and tickling children—gave very small percentages among the teachers one per cent and five per cent respectively. Asking silly questions was a poor third cause for teachers' laughter—15 per cent. This was a little higher, however, than 6th grade children's percentages of eight per cent and seven per cent Getting very angry was seven per cent for teachers, about one per cent for children The increasing funniness to children of playing tricks gave higher percentages, 11 per cent, for the older groups than the seven per cent indicated by the teachers Only three per cent of the teachers thought grown people falling down was funny compared with 10 per cent for Grade 3 boys, three per cent for Grade 6 boys, and about six per cent for the girls of all four grades

6 *People Look Funny*

Table 8 shows the percentages for what the children thought made

TABLE 8

PERCENTAGES OF CHOICES OF THINGS THAT MAKE PEOPLE LOOK FUNNY BY SEX AND GRADES

	Grades								
	Boys				Girls				Teachers
	3	4	5	6	3	4	5	6	
	a—Both sexes increasing								
Cross eyes	6	6	9	11	5	7	9	9	6
Wear funny hats	7	7	13	13	10	15	16	21	31
	b—Both sexes decreasing								
Have big ears	27	27	15	21	23	19	20	17	11
Shoot each other	10	10	7	4	6	4	3	1	0
	c—Both sexes little changes								
Very fat	21	21	22	19	23	22	21	20	14
Have big noses	19	19	18	16	18	15	19	16	8
Very thin	4	4	6	4	5	4	2	3	4
Fall down	2	2	2	5	2	2	2	6	10
Sing	2	2	2	2	5	3	1	2	12
Wear spectacles	1	1	2	4	2	7	3	4	2
Have no arms	2	2	2	0	2	2	0	0	0
Are in accident	0	0	0	0	0	0	4	0	0
Others									2

people look funny Having big ears and big noses, and being very fat seemed to have been the funniest things in this list of items The first two, and especially having big ears, apparently made less and less humorous appeal with age Being too fat, however, continued to be just about as uproarious for 6th grade as for 3rd grade children No sex differences regarding these points appeared to be present Another cause for laughing which seemed to decrease with age was shooting each other Both boys and girls of the 6th grade considered this of very little humor, the girls being more pronounced in this respect perhaps, than the boys Both sexes showed possibly a slight trend toward seeing more humor in cross-eyes as age increased Both sexes indicated marked increase in humor in seeing people who wear funny hats The girls, especially in Grade 6, gave somewhat greater percentages than the boys.

Being thin was nowhere near as funny as the opposite Probably being very fat and being very thin do not seem to be opposites to New York City children Possibly at such early ages "to slenderize" is already considered to be ideal. People falling down, singing, wearing spectacles, being armless, or in accidents were voted very low percentages.

The teachers' ideas about funny looking people are very interesting Needless to say, "wearing funny hats" lead all others by more than 2 to 1, having 31 per cent as compared with 21 per cent by the 6th grade girls who gave the highest percentage of all the children's groups on this matter. The teachers' second choice was being very fat. The percentage was less than that of any of the children's groups The teachers considered people who sing and fall down as funny looking in contrast to the very small percentage of the children's groups. The teachers gave fairly high percentages for big ears and noses and for cross eyes, but smaller percentages than the older children's groups. The possible increasing trend of the children in regard to cross eyes apparently must reverse itself somewhere beyond Grade 6 and return to about the 3rd grade figure which was the same as that of the teachers

7. *Things That Make Me Laugh*

Table 9 shows the figures for what things were thought funny Clowns lead the list with no sex differences apparent and with an odd duplication of percentages in Grades 4 to 6 inclusive. The high

TABLE 9
PERCENTAGES OF CHOICES ON THINGS THAT MAKE ME LAUGH BY SEX AND GRADES

	Grades								
	Boys				Girls				Teachers
	3	4	5	6	3	4	5	6	
a—Both sexes increasing									
Dummies	14	12	16	17	13	13	18	14	10
Queer words	1	2	5	5	1	2	5	8	20
b—Both sexes decreasing									
Jack-o'-lantern	6	7	5	4	7	9	6	3	2
c—Boys little change, girls decreasing									
Chinese	12	9	12	8	8	8	5	2	0
House with queer windows	8	7	8	4	11	10	9	7	4
d—Little change, both sexes									
Clowns	21	26	26	26	24	27	27	27	35
Boys in girls' clothing	12	9	9	12	14	11	14	20	10
Girls in boys' clothing	12	10	8	11	10	12	11	13	7
Acrobats	5	7	8	7	6	4	3	4	4
Cowboys	4	4	1	3	4	2	2	1	2
Cats	3	3	2	2	1	2	1	1	3
Chil jumping rope	1	3	1	1	1	0	0	1	0

percentages given by the children, 26 per cent by boys and 27 per cent by girls, will go still higher, however, if the teachers' percentage, 35 per cent, is a valid indication The children placed next to clowns the incongruous costume of boys' clothing on girls and girl's on boys Dummies also seemed funny to the children and increasingly so with children's ages. Compared with the teachers these percentages were high, which indicate that somewhere above Grade 6 a decline would bring this interest back to the smaller percentage shown by the teachers

The greatest contrast between teachers and children was on the item of queer words The interest of the children seemed to be beginning at Grade 3 and was rising quite sharply by Grade 6 Apparently that increase would continue as the 6th grade children's percentages were far below 20 per cent shown by the teachers Both boys and girls indicated decreasing appeal of jack-o-lanterns, the 6th graders dropping to practically the teachers' percentage of two. The most evident sex difference was about the Chinese The boys seemed to think them somewhat funnier than did the girls The latter also

indicated a rapid drop to almost zero by Grade 6, the figure for the teachers. The remaining three things—cowboys, cats, and children jumping rope—were not considered by any of the children's groups, nor by the teachers, as being much cause for laughter

8. *Stories*

A discussion of the replies to Questions 8, 9, and 10 has been made elsewhere.[2] Questions 11, 12, 14, and 16 are briefly discussed later in this paper.

9. *My Teacher Makes Me Laugh*

Table 10 shows the results on checking 10 items indicating what teachers do that makes children laugh. Four of these—the teacher

TABLE 10

Percentages of Choices on How My Teacher Makes Me Laugh

	Grades								
	Boys				Girls				Teachers
	3	4	5	6	3	4	5	6	
	a—Boys decreasing, girls little change								
Jokes	17	18	14	12	15	18	17	17	25
	b—Boys increasing, girls little change								
Sarcasm	3	3	6	8	4	2	4	5	2
	c—Little change, both sexes								
Tell funny stories	16	19	17	15	18	21	16	18	20
Read funny stories	16	19	16	12	16	19	15	13	13
Laughs herself	13	12	9	12	14	13	12	12	20
Makes queer motions	10	10	11	9	10	10	9	12	8
Gets angry	9	7	6	9	5	3	7	8	6
Forgets	7	4	8	9	7	5	6	6	3
Makes mistakes	5	4	10	8	5	6	9	7	2
Punishes others	5	4	3	6	5	2	4	3	1

telling funny stories, reading funny stories, telling jokes, and laughing herself—gave about the same percentages, about 12 per cent to 18 per cent. The remaining six items were somewhat lower and seemed to be in the following order of importance: makes queer motions, gets angry, forgets, makes mistakes, sarcasm, and punishes others.

Sex differences were indicated in only two respects and in each the trends were not great. Boys seemed to show decreasing interest

[2]Florence Brumbaugh Stories children think are funny *Elementary English Review*, October, 1939.

with age in the teachers' jokes, while the girls, regardless of age, seemed to think them equally and quite good Secondly, the older boys showed slightly more humorous interest in the teachers' sarcasm. No group of either sex rated this very high, but the 6th grade boys gave it eight per cent compared with three per cent by both 3rd and 4th graders

When the teachers answered the check list at this point they read the item. *"My professor makes me laugh when."* In general the teachers agreed with the children. Their indorsement of the higher rated children's choices was stronger, however, and this was at the expense of the lower ranking items. The jokes and the funny stories read or told, were distinctly the most appealing humorous endeavor of the professors

10 *Jokes*

A collection of jokes was made from the 800 children with whom the preliminary work of the study was done. The pupils were asked to write a joke. Table 11 gives the percentages who wrote jokes, riddles, and made no replies

TABLE 11
RESPONSES BY TYPES MADE BY 800 CHILDREN WHEN ASKED TO WRITE A JOKE

	Grades									
	Boys				Girls				Total	
	3	4	5	6	3	4	5	6	*B*	*G*
Jokes	19	29	38	54	28	28	26	24	35	27
Riddles	37	31	25	30	23	24	26	44	31	29
Total	56	60	63	84	51	50	52	68	66	56
No response	44	40	37	16	49	50	48	32	34	44

A rather striking difference is shown between the boys and the girls The percentages of boys writing jokes increased markedly grade by grade, while the percentages for girls changed little if at all Furthermore in Grade 3 more girls than boys wrote jokes. Perhaps the superior penmanship and language abilities usually found among girls of the third grade would largely explain that In Grade 5, however, half again as many boys as girls wrote jokes, and in the sixth grade more than twice as many

The sex differences in the percentages writing riddles for jokes are not as distinct There seems to have been a slight trend to de-

crease by the boys as age increased. The girls showed little difference until Grade 6, where a large increase appeared Just what that signifies is doubtful, because of the relatively small number of cases involved.

The sex differences in the percentages making no responses are, as in the case of the number writing jokes, much more suggestive Fewer and fewer boys, by grades, failed to respond with something In the case of the girls, however, nearly half of the 3rd, 4th, and 5th graders made no response The drop in no responses for the 6th grade girls was proportional to the large jump in the riddles given by the girls in that grade.

Question XIII of the check list was *"If you tell jokes, where do you get them?"* The data on this question were taken from the replies of the 800 children who marked the check list It will be remembered that these were different children than those who wrote the jokes just discussed above. Six sources were suggested on the check list (found from the preliminary work with the children on the questionnaire) and the children were told to add any other sources they might have had Only 16 children gave other sources. Table 12 shows the results for the answer to this question. Since varying numbers of answers were made by different children, the number of times each answer was indicated is given, instead of percentages

TABLE 12

NUMBER OF TIMES EACH SOURCE OF JOKES WAS INDICATED BY CHILDREN WHO ANSWERED THE CHECK LIST

	Grades								
	Boys				Girls				Total
	3	4	5	6	3	4	5	6	*B* & *G*
Joke book	59	72	52	64	66	73	42	57	485
Radio	54	59	48	71	61	50	43	66	452
Other children	38	35	45	57	55	42	38	71	381
Parents at home	16	21	14	24	28	23	13	30	169
Newspapers	35	24	13	32	30	7	6	19	166
Magazines	15	25	16	25	28	20	17	18	164
Miscellaneous	1	3	2	2	1	2	3	2	16
Totals by grades	218	239	190	275	269	217	162	263	1833
Totals by sex		922				911			
Total grades	3 and 4								913
Total grades	5 and 6								890
Average per pupil									2 8

A total of 1833 answers was made by the 800 children, or an average of 2 8 per child This seems to indicate that the children questioned were interested in jokes and put forth considerable effort to get hold of them The leading sources were about equally popular and evidently were very much sought for More than half of the children reported that they used joke books and the radio to get jokes Grade and sex differences were not marked, the lower numbers for the 5th grade girls being explainable, perhaps, by chance The third main source, other children, was not far below the first two In all grades this source was given by large numbers of both sexes. The numbers of children using the next three sources—people at home, newspapers and magazines—were much lower than the three leading ones, and each totalled practically the same for all groups together—164, 166, and 169 No marked sex or grade differences on these points were indicated

The totals by sexes and grades bear out the statement made above that there were no conspicuous differences among groups Similarly the totals of the two lower and the two higher grade groups, 943 and 890, show no significant differences. Apparently these children were interested in getting jokes regardless of grade or sex and drew heavily on out-of-school sources for them Teachers, school books, and other school materials seem to have been completely out of things as far as contributing to this sort of laughter was concerned

II. *Poems*

Giving the titles of poems presented more difficulties to the pupils than any other part of either the questionnaire or the check list. Twenty-one per cent of the boys and 18 per cent of the girls did not reply to this request in the check list Of those who did, many obviously wrote the name of any poem remembered, and not those that were humorous

Few of the poems had been heard outside the school, according to the answers made by the children This proved to be true, when the titles were checked. It was found that every poem that was selected by more than one pupil could be found either in the required literature for the grade, for which special texts were found, or in a basal reader used in the grades

Many titles were those of poems that had been memorized or studied in previous grades, but the most popular in every grade were

those assigned for that particular class. Nursery rhymes were given in every grade, including the sixth, and in almost equal numbers. The two favorites of the 17 rhymes were *Jack and Jill* and *Humpty Dumpty*.

Each grade had two or more favorites liked by both the girls and the boys. The third grade's first choice was *The Owl and the Pussy Cat* (Lear), which was almost as popular in the fourth grade, and less so in the upper grades. Nursery rhymes were next in importance in grade three, followed by *Popeye in Verse*.

The fourth grade chose *Mr. Nobody* (anonymous), found in the required literature as well as in a basal reader. *The Mountain and the Squirrel* (Emerson) and *Book Houses* (R. L. Stevenson) were second choices and were in readers used by the fourth grade.

The fifth grade liked *A Tragic Story* (Thackeray), *The Twins* (Hunt), and *The House with Nobody in It* (unknown). It was in this grade that the only limerick was mentioned, *The Lady and the Tiger*. All of these were included in readers used in the grades.

Father William (Lewis Carroll) and *The Curious Case of Ah Top* were the sixth grade's choices. In both the fifth and sixth grades, the boys and the girls showed great differences in their choices of poems. All others than those poems named were either girls' or boys' preferences only.

Poems by R. L. Stevenson were better liked than those of any other author, although it is difficult to see why some of them would be considered funny.

A very few street chants were mentioned,

"No more pencils
No more books
No more teachers'
Dirty looks."

The total numbers of poems of different titles for each grade were as follows: Grade 3, 56; Grade 4, 59; Grade 5, 52; Grade 6, 86.

12. *Miscellaneous*

The data for the following proved inadequate for worthwhile analysis: Question 12, Funny songs; and Question 14, Do you laugh often, sometimes, almost never?

The analysis of Question 16—*Did you think this test was Funny?* gave overwhelming approval (Table 13).

TABLE 13

	Grades								
	Boys				Girls				Total
	3	4	5	6	3	4	5	6	
Yes	84	98	89	95	92	94	65	100	60
No	16	2	8	4	3	2	2	0	30
No response	0	0	3	1	5	4	33	0	10

C Summary

Table 14 gives a selection of 20 things indicated by the children as the funniest one, two or three items on each question of the check

TABLE 14

Summary Table of Things Thought the Funniest by Boys and Girls in Grades 3 and 6

	Grade 3		Grade 6	
	Boys	Girls	Boys	Girls
a—Little or no change				
Charlie McCarthy	18	23	21	22
Very fat	21	23	19	20
Make faces	16	18	18	20
Teacher tell funny stories	16	18	15	18
Teacher tell jokes	17	15	12	17
Dunces	14	13	17	14
b—Decreasing Funniness				
Children tickle	30	30	17	24
Big ears	27	23	21	17
Grown people tickle	21	22	13	17
Donald Duck (comics)	21	13	9	5
Teacher read funny stories	16	16	12	13
Big noses	19	18	16	16
Grown people kiss	24	22	13	8
c—Increasing Funniness				
Clowns	21	24	26	27
Children tell jokes	17	17	29	27
Katzenjammer Twins	17	13	22	25
People wear funny hats	7	10	13	21
Laurel & Hardy	7	6	22	18
Ritz Brothers	10	10	19	18
Burns & Allen	9	11	12	19

list In order to simplify the table only Grades 3 and 6 have been shown, but the sexes have been kept separate. The arrangement is roughly in order of the size of the percentages, first for those

items showing *little or no change*, then for those showing a decrease, and finally for those showing an increase from Grades 3 to 6

The table has limited validity as a *presentation of comparative* funniness of the items, because no such comparison was made by the children All they were asked to do was to compare the items under each question as they went along on the check list Thus they selected the funniest movie characters, then the funniest radio people, *etc.* They did not compare the funniest movie characters with the funniest radio ones, and so forth The summary table, therefore, may not be interpreted that way, either If all the items shown in the table had been compared with one another the percentages would, of course, have been much different for most of them and perhaps would have been very different, relatively, to one another.

The table, however, brings together the items which the larger percentages of the children thought were especially funny. It also shows the marked trends for these items in Grades 3 and 6 An analysis may be made of the funny things which show similar trends, and there seems to be a limited degree of similarity that may be characteristic of the items in these groups The trends are clearest in the last two parts of the table All of the items in the group showing decreasing funniness, except the teacher reading funny stories, are things which might be expected to be relatively less mirth provoking to children as they grow older—being tickled, seeing people with big ears and noses, the fanciful Donald Duck, and grown people kissing each other. Most of these things (perhaps not the last one) are simple and compared with other interests have little of the deeper meanings which experience teaches. On the other hand, the things showing increasing funniness seem, in most cases, to have more meaningfulness in them—jokes, the movie and radio comedians, clowns (the comedians are supposed to be artists in mirth provoking, and art demands meaning) and the Katzenjammer Twins, who typify experiences children would like to have One exception (in the opinion of one of the writers) is found in this group: people who wear funny hats. Perhaps in a less obvious way that, too, becomes more meaningful with age, for example as to incongruity.

The first group of items in the table, things which show little or no change, seems to be much more of a mixture of simple and complex reactions than either of the others. For example, Charlie McCarthy has a wealth of fun-making meaning, while being very fat and making faces seem relatively simple

These suggestions are made with the certainty that they are at best only part of the explanations for the funniness seen in the various items according to children's ideas about them. There is little doubt that many undetermined factors were operating in the responses of these children, and even in the case of the simplest and most general incitors to laughter, probably different factors appealed to the risibles of different children.

Department of Education
Hunter College
695 Park Avenue
New York City

The Journal of Genetic Psychology, 1940, **57**, 31-47

CONTRIBUTIONS CONCERNING MENTAL INHERITANCE II TEMPERAMENT*

Department of Psychology, Clark University

Raymond B Cattell and E Virginia Molteno

A. Limitation of Aim of Investigation

The present contribution is not concerned with defining temperament, or surveying past general evidence, but sets out to achieve new experimental evidence in one particular field

Temperament "types" are as numerous as they are ill founded Experimentally founded temperament factors, vectors, or "gestalten," which have received statistical confirmation from factor analysis and in the realm of the normal psychology of personality are, on the other hand, few Indeed, general emotionality (2), surgency (3, 19, 20) (the unchanging core in "extraversion"—underlying particularly the traits cheerful, sociable, quick of apprehension, impulsive, humorous, informal, etc.), and possibly the pattern of characteristics concomitantly varying with perseveration (8, 15) are practically the only factors on which the objective psychologist can yet depend

Wishing to deal with the problem of inheritance in a quantitative manner, and to produce results which could be integrated with further inquiries, we limited our study of temperament to factor patterns, and in fact to those which already admit of fairly valid measurement by laboratory tests

Surgency, as based on estimates, has been repeatedly shown to correlate with the test factor "fluency of association" (7, 8, 20) to the extent of 0 6, (ignoring correction for attenuation or functional fluctuation) while the consistency of fluency tests is the same as that of intelligence tests of the same duration, namely about 0 9. The surgent temperament can therefore be measured with more reliability at present than is possible with other temperament patterns or tests

Perseveration is not so satisfactory Though its existence as a general factor is shown by all researches which use adequate precautions, the consistency of measurement is seldom high, partly on ac-

*Received in the Editorial Office on May 6, 1939

Diff. adopted sibs.

count of the brevity of tests, necessitated by fatigue. But its importance for personality, normal and abnormal, has been indicated by so many researches (4, 5, 7, 15, 17) that its present obscurities cannot negate its use.

In investigating the interplay of inheritance and environment we decided to follow the method of comparing identical twins with fraternal twins of the same sex, i.e., monozygotic and dizygotic individuals reared in the same environment. Such a decision does not imply that we consider this method the only sound one in heredity studies, or even the best, but rather that it happens to be practicable for our particular inquiry and permits our results to be directly compared with others which have so far kept to this method. The fashionableness of the method should not blind us to the fact that it achieves a comparison of the effectiveness of environment and of heredity only in one particular situation: that concerned with variations of family environment reacting with intra-familial variations of heredity.

The study of siblings reared apart, or of children and parents who happen to have been separated, would, on the other hand, provide quite as valuable evidence of a direct nature. And this would be only a shade less desirable, because only a little less applicable to the solution of social and personality problems, than that from the ideal but scarcely attainable study of an adequate group of identical twins reared apart. Indeed, in view of the possible existence of a special and systematic error in the method of comparing monozygotic and dizygotic twins, which is discussed later, we would suggest that investigators would do better to study siblings in comparison with unrelated adopted children reared in the same family. Such instances would be only slightly more difficult to discover than suitable twin pairs and would permit the calculation of a ratio[1] expressing hereditary effectiveness in producing resemblance within the family. This would be more directly applicable than other ratios to many calculations attempting to weigh the influence of family heredity in individual and social situations. The first of the above-mentioned methods would also lead, as the twin method does not, to assessments of the relative potency of inter-familial rather than intra-familial variations in heredity and environment.

---

[1] $R = \dfrac{\textit{Diff}\text{ random children}-\textit{diff}\text{ adopt sibs}}{\textit{Diff}\text{ adopt sibs.}-\textit{diff}\text{ true sibs.}}$

Some indications that perseveration may have an hereditary basis are already known to most research workers in the field. Wynn Jones (12) found an inter-sibling correlation of +0.3 in a sufficiently large group of siblings. Rangachar (16) and one of the present writers (5) have independently shown that there are significant racial differences in mean perseveration score. Finally E. P. Yule (21) in a very thorough twin study of which the present research is a repetition, except for some changes in the "*p*" test battery and a restriction of age range to avoid spurious correlation, has shown definite evidence of inheritance. On the other hand, it is known that the extreme perseveration scores of certain psychotics tend to return to normal as the individuals improve in adjustment (17), and that among normal subjects perseveration score varies with emotional adjustment (7) and with fatigue (6).

Nothing whatever is known, however, about the relative importance of nature and nurture in determining the individual's temperamental balance with regard to surgency or "fluency of association." From the fact that manic-depressive psychosis—which may be regarded as a tendency to extreme manifestations of surgency and desurgency [since the patterns agree (20)]—is known to have a definite hereditary basis, one may reason by intrapolation that among normals the tendency to high variability of surgency is inherited, and indeed there is some slight direct evidence that high variability is a characteristic of certain individuals making extreme scores (7). Apart from these indirect indications, however, our study of *fluency of association*, unlike that of *perseveration*, breaks entirely new ground.

B. Account of Methods

Eighty-nine boy twins, restricted to 12 or 13 years of age, were taken for the experiment from the elementary schools of London, Leicester, and Derby (England). The selection was purely random, except that the intelligence range and social background were those of the elementary school from which scholarship children have already been promoted. Among the Leicester cases, for added interest, four problem children, who happened to be twins in this age range, were included from the school psychological clinic.

The sorting into identical (monozygotic) and unlike (dizygotic) twin pairs was made on the basis of the following physical characteristics:

Agreement in: (1) Height
(2) Head circumference
(3) Eye color
(4) Hair color
(5) Hair texture
(6) Hair whorl direction
(7) Skin color
(8) Skin texture
(9) Size of feet
(10) Size of hands
(11) Arrangement of teeth
(12) Shape of ears
(13) Handedness
(14) Appearance of front face
(15) Appearance of facial profile
(16) Fingerprint patterns

The experience of the principal earlier investigators (9, 14, 19) of physical twin resemblances was our chief guide in this sorting, and in the instances where any doubt whatsoever existed as to the correct category of the twins, we obtained an independent judgment. For this we are indebted to the Galton Laboratory When eye color, skin texture, hair, and bodily proportions are alike, but some suspicion of dizygosity still presented itself, the final decision was made with the aid of finger prints, according to the techniques of Newman (14) and Stocks (19). In order that twins should be considered monozygotic the cross agreement between the hand of one twin and either hand of the other twin had to be greater than that between the hands of the same twin. Furthermore, this agreement had to be found on not fewer than seven of the ten digits.

It is perhaps too readily overlooked in researches using the twin comparison method of heredity study that even today the above criteria do not permit an infallible division of twins into two types Consequently, as an examination of the formulae below will show, the usual estimates of the variance due respectively to heredity and environment are likely to err by giving too much weight to environment and in proportion to the amount of adulteration of the twin groups with twins of the opposite type To avoid this as far as possible in the recent research, five pairs of twins were rejected because of doubtful diagnosis, leaving 53 twin pairs diagnosed as dizygotic (fraternal) and 31 as monozygotic (identical)

The "*F*" tests by which the twins were tested were given under the standard conditions and with the method of scoring normally used in clinical work and described elsewhere (18). The "*P*" tests were selected from two standard batteries published elsewhere (8, 18), on grounds of high "*P*" saturation and small demand on intelligence The battery contained the following tests Each symbol followed by a comma indicates that one row of such symbols was made by the subject during a period of 15 seconds, with a rest pause before the next row. Further details of administration are given in the place cited.

PERSEVERATION TESTS

1 √, √, √, √, √,
2 H, ⌶, H, ⌶, H ⌶, H ⌶, H ⌶, H ⌶,
3. 234 (*normally* written), 234 (written with *backward* stroke of pen), 234 (n), 234 (b), then alternatively normal and backwards
4 W, W, (normally written), W, (backwards), alternating
5 abc, ABC, abc, ABC, aAbBcC, aAbBcC, aAbBcC, aAbBcC
6 △, ▽, △, ▽, △ ▽, △ ▽,
7 Reading off color spots on a sheet. "Red, blue, etc "
Normally, Normally, Reversed naming, Normally, Reversed naming

The "*F*" test, which consisted of (*a*) additions to pictures, (*b*) word lists, (*c*) form completion, (*d*) story completion, (*e*) ink blot interpretations, required 15 minutes of actual testing time and had a consistency coefficient of 0 84 (split-half consistency 0 91). The "*P*" tests depended on 11 25 minutes of actual testing. Their consistency, over a week's interval, proved to be low, being only 0 21 in a group test situation, but rising to 0 51 in the present individual test situation.[2] Throughout the tests were given individually, the "*F*" test following on the "*P*" test They were administered as far as possible at the same time of the day, to obviate or diminish false individual differences arising from fatigue differences (5) The "*P*" tests, as will be seen in the handbook cited, were scored by the $\frac{X}{Y}$ formula (X being the original activity and Y the impeded activity) to reduce or eliminate effects due to variability of speed of writing, and the score on the battery was taken as the sum of the $\frac{X}{Y}$ quotients for all seven of the sub-tests

[2]It was possible to retest only part of the group

C. Analysis of Results, with Regard to "Fluency"

Both with fluency and with perseveration the relative similarity of monozygotic and dizygotic twins can be conveniently expressed in either of two mathematical forms: (*a*) on the basis of the mean difference between twins, (*b*) on the basis of the correlation coefficient between twins.

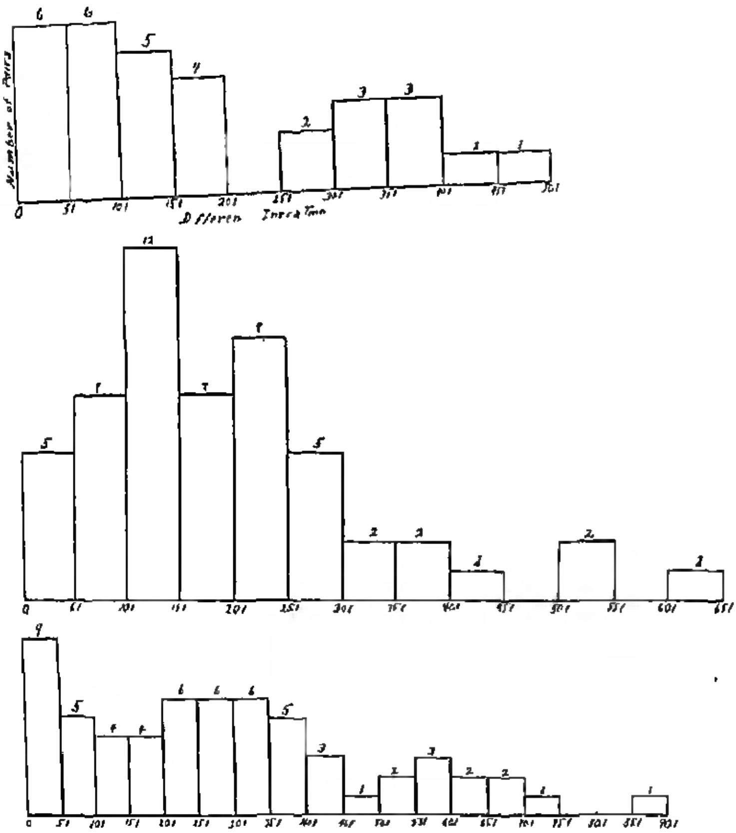

FIGURE 1

Above: for 31 pairs of monozygotic twins
Middle: for 53 pairs of dizygotic twins
Below: for 120 individuals taken at random in 60 pairs

The actual intra-twin difference in "F" score for monozygotics and dizygotics were distributed as shown in Figure 1

The means of the above distributions are shown in Table 1 The

TABLE 1

"F" Differences, Intra-Twin Pair

		1st PE		2nd PE
Monozygotic twins	17 4	±1 64	or	±1 80
Dizygotic twins	19 6	±1 20	or	±1 49
Individuals at random	28 7	±0 91	or	±1 02

above probable errors have been calculated on two distinct bases The first is the usual PE of a mean, which is not accurately applicable to the present measures because the distribution is badly skewed, indeed since each twin difference is entered once only, in a positive direction, the underlying polygon may be considered as representing one half of a normal distribution curve. The second PE considers each difference of a twin from his fellows as a deviation, which is entered up twice (as in the double entry method in correlations), once in a negative and once in a positive direction The probable error is then calculated as the probable error of a mean deviation The present writers consider the second PE as the more accurate one, but have inserted the more usual PE to show that the conclusions are not altered by its use.

Clearly, from Table 1, twins resemble each other far more than randomly chosen individuals in the general population On the other hand, the difference between the resemblance of mono- and dizygotic twins does not appear to be statistically significant, as indicated in Table 2.

TABLE 2

		PE of diff 1		PE of diff 2
Diff of Mono- & Dizygotic twin differences	= 2 2	±2 03	or	±2 32
Diff of Monozygotic & random individual differences	= 11 3	±1 87	or	±2 08
Diff of Dizygotic twin and random indiv diff	= 9 1	±1 51	or	±1 81

On the other hand, when the frequency distributions (of the intra-twin differences) are themselves compared for agreement by goodness of fit, there is found to be a significant difference between

mono- and dizygotic twins, for chi square works out at 9.80, giving a P value of 0 05, i e , there is a twenty to one likelihood that the divergences between the two distributions are not due to chance.

We shall therefore proceed to examine the mean difference between mono- and dizygotic twins to see what light it throws on the parts played by heredity and environment in *determining the individual's* fluency Various formulae have been applied to twin study data, of which the simpler, e g ,

$$\frac{\textit{mean difference of dizygotics}}{\textit{mean difference of monozygotics}}$$

fails to give any adequate or clear expression of heredity-environment ratios We shall apply the more satisfactory formula used by Hogben (11) and others

$$T = \frac{\textit{Mean diff of dizygotics} - \textit{Mean diff. of monozygotics}}{\textit{Mean difference of monozygotics}}$$

This gives a ratio expressing the importance of heredity, relative to that of environment, in determining the mean differences between members of the same family, on the assumption that there is no difference between the environments of mono- and dizygotic twins, and that their environments as individuals differ from each other as much as the environments of other members of the family (siblings) These assumptions can be questioned and constitute weaknesses setting limitations to the usefulness of the twin-study method If for the moment the above indication of a possibly greater mean difference between dizygotic and monozygotic twins is accepted at its face value, T is found to have the value 0 126, suggesting that differences within the family environment are about eight times as im-

TABLE 3

Intra-pair "F" correlations*	
For Monozygotics	$r_m = 0\,656 \pm\ 048$
For Dizygotics	$r_d = 0\,594 \pm\ 042$

*These and all subsequent correlations are calculated by the method of double entry (each twin being entered once as an abscissa and once as an ordinate in the regression diagram) usual in twin studies in which there is no systematic ground for considering *one of each pair to belong to a* certain class

portant as hereditary segregation of genes in accounting for individual differences in fluency.

Applying now the approach by correlation methods we find the coefficients to be as indicated in Table 3.

Corrected for test attenuation these figures become respectively 0.78 and 0.70, indicating again a very considerable twin agreement in contrast to the lack of agreement naturally found in the population at large. But here too the difference in favor of greater resemblance of monozygotics cannot be accepted as more than an indication, for the difference is no greater than its probable error (0.064). Again, if it were taken at its face value and employed in the formula $t^2 = \frac{r_m - r_d}{1 - r_m}$, which expresses the variance due to heredity as a fraction of that due to environment, we obtain a value for t^2 of 0.179, indicating that environment plays a larger part than heredity in determining fluency variations within the family.

D. Analysis of Results, with Regard to Perseveration

After examining the distribution of intra-twin pair perseveration differences, as shown in Figure 2, one is not surprised to find on calculation that the resemblance of monozygotic twins, as shown in the mean difference, has no significant difference from that of dizygotic twins. Indeed, the difference of monozygotics is slightly greater than that of dizygotics, the figures being as indicated in Table 4.

TABLE 4

"P" Differences, Intra-Twin Pair

Monozygotic twins	=	0.8635	±0.085	or	±0.088
Dizygotic twins	=	0.6603	±0.048	or	±0.049
Individuals at random	=	0.7379	±0.025	or	±0.036

There is no evidence of any hereditary factor in perseveration; indeed, the difference, such as it is, is in the wrong direction for such an influence to be deduced. The difference between dizygotics and individuals unrelated is also scarcely statistically significant, being 0.776±0.54 or ±0.61. An examination of the frequency distributions of mono- and dizygotic twin differences by means of chi square again reveals no significant divergence between the two types of twins.

By the correlation approach, however, we encounter positive evi-

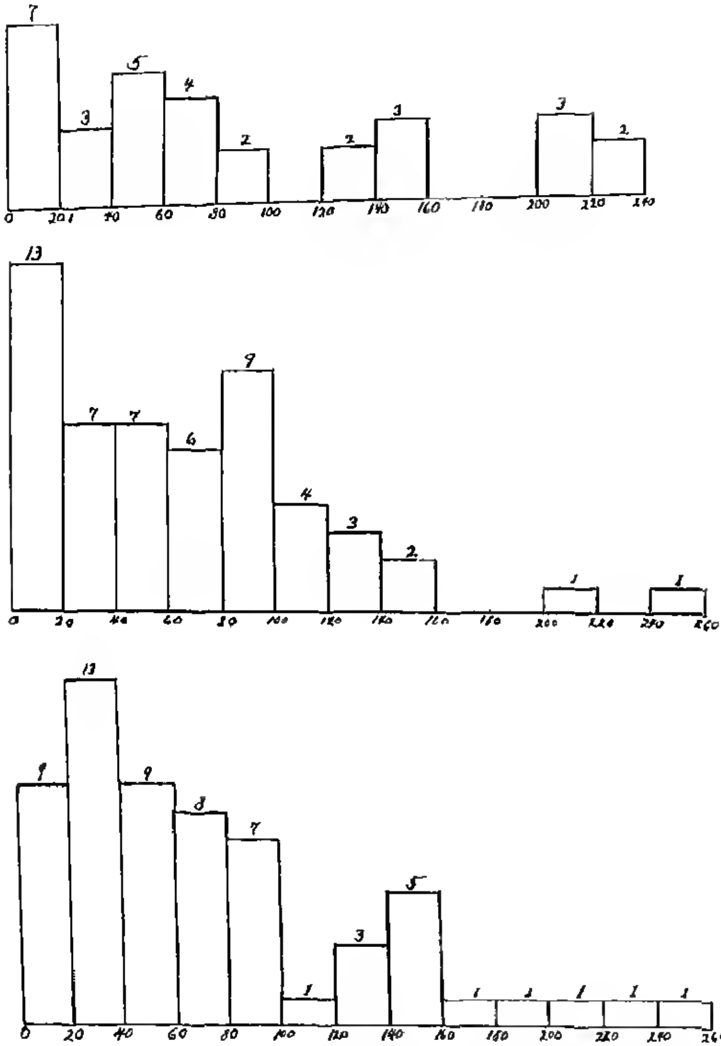

FIGURE 2

DISTRIBUTION OF "F" DIFFERENCES

Above for 31 pairs of monozygotic twins
Middle for 53 pairs of dizygotic twins
Below for 120 individuals taken at random in 60 pairs

dence that the larger group of twins, at least, shows greater perseveration resemblance than exists in the general population

TABLE 5
"P" CORRELATIONS, INTRA-TWIN PAIR

For Monozygotics	$r = 0.101 \pm 0.085$
For Dizygotics	$r = 0.465 \pm 0.051$

E Discussion of Anomalies in Relation to the Twin-Study Method

Two unexpected divergences from previous notions occur in the above observations First, in contrast to Yule's painstaking research described below, we find no hereditary influence in perseveration Second, the distributions of differences of monozygotic twins, both in fluency and perseveration, but particularly in the latter, show a small but excessive number of instances of very great twin difference These few highly divergent cases are alone sufficient to account for the monozygotic twin correlation in perseveration being actually inferior to that of dizygotic twins

Inspection of the distributions of Figure 1 almost suggests a bimodality of the monozygotic group distribution Calculation, however, fails to confirm this suspicion positively, chi square having a value of 2 865, giving $P = 0.25$ Similarly a test for kurtosis shows that one sample in five from a normal distribution would have as much kurtosis

Nevertheless, there is at least a tendency in the appearance of the distribution diagram for monozygotics to fall into two types, a main group and a small number of nonconformists, which justifies our questioning whether the calculation of nature-nurture ratios on the assumption of a uniform monozygotic twin type is sound.

In the first place, psychoanalysts and clinical psychologists have suggested a "protest" reaction by which one or both of the twins react away from the personality type of the companion by reason of the need, felt by an individual conscious of his individuality, to distinguish himself emphatically from another being who resembles him too closely Lange's study (13) of criminal twins contains several striking instances of this semi-conscious exaggeration of whatever slight differences have been brought into being by the chance influences of environment

The opposing view, that monozygotic twins accept their resemblance and grow to resemble each other more when reared in mutual contact than when apart, is, of course, the commonly accepted one· von Bracken (1) has stressed this in regard to appearance, and Gottschaldt's recent survey (10) claims the same for mental resemblances.

The protest reaction may be rare, but it undoubtedly exists, and the present writers have observed in the clinic definite instances in which the chance development (through illness) of certain traits in one twin has forced upon the other precisely opposite or complementary characteristics

Possible biological causes of anomalous monozygotic twins have also to be considered, though they are not so convincing.[3] Generally they could scarcely account for the differences we are considering, because they would probably lead to the twins' being classified as dizygotic on physical grounds, from the beginning.

A method of testing the "urge to divergence" hypothesis by the present data suggests itself Most evidence (5, 7, 12, 16) indicates that there is a certain stability about the perseveration index of every individual—a certain level congenitally or environmentally acquired which has become most natural to him and about which his score tends to oscillate with different states of fatigue and emotional adjustment. If now the individual forces himself to abandon that particular adjustment in order to destroy a too closely sympathetic responsiveness to the perseveration state of the other twin, it is reasonable to expect that distortion would occur in other features of personality structurally dependent upon the first

Now one of the present writers has observed in cases reported elsewhere (5, 7) that, as judged from a clinical standpoint, increased activity of the will and increased strain, which is typically accom-

[3]It is considered possible, though extremely rare, for a type of twin to arise midway between mono- and dizygotic forms, due to the fertilization of a divided ovum by two distinct spermatozoa

The writers are indebted to Dr. L S Penrose for calling their attention to another possibility suggested by Dahlberg (*Twin Births and Twins from an Hereditary Point of View*, Stockholm, 1926), namely, that distinctly unlike monozygotic twins may be produced by the vertical division, at the two or four cell stage of multiplication of an embro originally destined to produce a strongly asymmetric individual. Right-left asymmetry seems to be established at a very early stage of cell division, but there remain many difficulties in this notion, one of which is the rarity of such strong asymmetry as would lead to the twin differences here under observation

panied by a reduction of perseveration index, frequently seems to result in a diminution of the individual's natural surgency and fluency of association If there is any tendency for the individual to adopt constitutionally a certain level of perseveration we should expect therefore, on this line of reasoning, that the member of the twin pair who reduced his perseveration markedly from the more normal figure of his fellow twin would also necessarily distort, among other things, his fluency score in the same direction

This hypothesis can readily be put to the test, for it implies that among identical twins "P" and "F" intra-twin differences are correlated and probably in a positive sense Correlating "F" and "P" differences of the same twins without regard to the sign of the difference we obtain 0 342±0 109, and, when the sign is taken into account, 0 165±0 117 For dizygotic twins, on the other hand, the figures are respectively 0 113±0 092 and —0.052±0.093 Evidently as our hypothesis requires, the tendency to correlated divergences is something which appears among monozygotic and not among dizygotic twins,[4] though the divergence of "P" in one direction does not consistently result in a change of "F" in the same direction It should be pointed out that the actual "P" and "F" scores do not themselves correlate, either for individuals in the monozygotic pairs (0 14±0 13) or the dizygotic pairs (0 08±0 09)

This finding that "P" and "F" differences are correlated among monozygotics, the twins who differ most in "P" tending to differ most in "F," supports the inference that some sort of "protest" reaction is occurring strongly in certain instances. As we have seen above, if a few anomalous and highly divergent monozygotic twin pairs were omitted, the "P" score records would show greater than that found among dizygotics. In short, our negative result with regard to hereditary influences in perseveration cannot be set in the balance to nullify Yule's positive finding. If accepted, however, this line of argument implies that the use of monozygotic twins in heredity studies, on the assumption that the members of each pair enjoy the same environment to the same extent as dizygotics, cannot be justified

Furthermore this hypothesis throws a disturbing light on the

[4]It is worthy of note that the finding with these factors stands in contrast to the general rule pointed out by Thorndike, that the differences of twins in various characteristics do not positively intercorrelate

commonly accepted facts and theories regarding the independence of "*P*" and "*F*" as general factors For, if in any given individual an enforced modification of perseveration results in a corresponding change in fluency, it *follows that in the general population* some correlation of "*P*" and "*F*" scores might be expected It is possible that such a correlation, of small magnitude, does exist; no examination of a sufficiently large sample of the population has yet been published which would enable one to judge whether the small correlations sometimes found between "*P*" and "*F*" are significant[5] But since the correlation is certainly not of appreciable size we must *conclude that this* coordinated change of "*P*" and "*F*" is sufficiently rare, at least in the marked form observed in identical twins, to be swamped by other environmental and hereditary influences determining the variance of "*P*" and "*F*", or that it occurs with equal frequency in consonant and opposed directions

F Summary of Conclusions

1 In "fluency of association" ("*F*" factor) members of twin pairs show greater similarity than is found among pairs of the general population taken at random. In "perseveration" ("*P*" factor) the resemblance of twins is less marked, owing to there being no significant resemblance in the identical twins (monozygotic) group[6]

2 In "fluency of association" both the mean intra-twin differences and the correlation coefficients agree in showing slightly greater *resemblance of monozygotic than dizygotic (fraternal or unlike)* twins of the same sex Calculations on the basis of these figures show that environment is about eight times as efficacious as heredity in the production of the mean differences between children of the same family and five times as important in accounting for variance within the family These conclusions regarding "*F*" may presumably be

---

[5]The correlation repeatedly found is small and *negative*

[6]For the benefit of non-technical readers considering first these conclusions apart from the main article, some definition of these terms is desirable Dizygotic twins develop from two distinct fertilized ova The individuals resemble each other no more than ordinary brothers and sisters and are therefore *sometimes called "fraternal" twins They may be of like or* unlike sex Monozygotic twins arise from the splitting of a single fertilized ovum into two cells which proceed to develop independently They are rarer than dizygotic twins, are of the same sex and resemble each other so closely that they are also called "identical" twins

carried over directly to "*C*" or surgency[7] of temperament, since it would seem likely that the incompleteness of correlation between "*F*" and "*C*" is due only to functional fluctuation.

Incidentally this discovery of the importance of environment in determining the individual's "*F*" endowment lends support to the first of the two hypotheses suggested elsewhere, regarding the nature of fluency (or surgency), namely, that it is a function of inhibition (actually of deficiency of inhibition), rather than of energy endowment (5, 7).

3 There is no direct evidence in the above results that perseveration is anything but a function of the individual's past experience and environment Since this conclusion contradicts the trend of some previous indirect evidence by one of the present writers (5) and Rangachar (16) and the direct evidence of Yule's research (21), the conflicting evidence requires further examination and extension The presence of significant mean perseveration differences between physically distinguishable races in the same culture (5, 16) could, if necessary, be explained by less simple hypotheses than hereditary causation, notably by connecting the difference with environmental differences of social status or subtle traditional influences. On the other hand, Yule's conclusion could be doubted because no positive evidence of hereditary differences appeared when her perseveration tests were scored in the usual way, as here But it seems more reasonable to look for influences in our own research which might be screening the influence of heredity. Such an influence is described above

4 Among monozygotic, but not among dizygotic twins, the pairs which differ most in perseveration also tend to differ most in fluency of association This is compatible with the assumption of a specific divergent "protest" reaction of identical twins to the unusual environment of their identicality If "protest" is substantiated, even as a relatively rare occurrence among identical twins, it vitiates twin comparison technique in mental heredity study, for the greater environmental divergence of monozygotic twins hides to an indefinite extent the effectiveness of heredity. Even when these aberrant special "protest" cases are eliminated, however, heredity does not have as great an influence as environment in determining either perseveration or fluency status

[7]For a description of perseveration, fluency and surgency, as general temperament factors, see (8)

The Journal of Genetic Psychology, 1940, 57, 49-55

THE INTELLIGENCE OF FOSTER CHILDREN*[1]

Child Guidance Service of Sangamon County, Springfield, Illinois

George S. Speer

There have recently appeared a number of studies (2, 3, 4, 5, 7, 8) from the University of Iowa purporting to indicate a gain in *IQ* when the child is placed in varying stimulating environments: nursery school, adoptive homes, and so on. Conversely, these studies have indicated that children who are in a non-stimulating environment (their own homes, or a poor orphanage) are reduced to the level of the feebleminded. Simpson (1) in a recent critical analysis of these studies has concluded that these results are largely, if not entirely, due to "the pitiable inadequacy and shiftiness of the statistical analysis, and of the utter lack of dependable evidence in support of the glamorous claims which are made."

The validity of the Iowa claims are important to child-placing agencies. If these claims are correct, they should have an important influence on child placement policies, as well as on the educational procedures of school systems. If, on the other hand, these claims are not substantiated by other investigators, this information should be published before unwise foster home placements are made on the strength of the Iowa claims.

As a part of a comprehensive study of dependent children placed in foster homes by the Children's Service League, the child placing agency in Sangamon County, we have been interested in the mental development, as indicated by Stanford-Binet *IQ*, of these children in their own homes prior to placement, and development after placement in foster homes.

All of these children came from economically underprivileged homes. They have been declared dependent by the Juvenile Court, guardianship transferred to the Children's Service League, and placed in foster homes selected and supervised by the Children's Service

*Received in the Editorial Office on May 20, 1939.

[1]Read at the Symposium on the Constancy of the *IQ* at the 14th Annual Meeting of the Midwestern Psychological Association, Lincoln, Nebraska, May 6, 1939.

League.[2] The age at which they have been placed has not been affected by any factor other than need. When the parents died, deserted, or were declared incompetent, all of the children under 16 years were declared dependent, and placement and supervision undertaken by the Children's Service League.

The homes from which these children came were of a very low economic level. In addition, many social factors combined to make these homes undesirable environments. The parents were, in many cases, emotionally unstable, alcoholic, irresponsible, delinquent, unconcerned about regular school attendance for the children, uneducated, and, in general, unaware of or unresponsive to those duties which we normally consider to belong to parenthood.

A random sampling of 15 true homes shows extreme poverty in every case; eight of the parents are mentally defective, as well as a number of the collateral relatives; eight were sexually degenerate and four were known to be venereally diseased; in three instances, the father had had incestuous relations with his daughters; six of the parents were insane and in state hospitals and, in addition, at least two of the immediate relatives were also insane; four of the parents were alcoholic; four were dead; three had deserted their families.

There were 184 children in the group which we have studied. The mothers of 68 of these children are definitely mentally defective (median *IQ*, 49.0) and are all confined in state institutions.[3] The mothers of 57 children are apparently normal by social criteria, but have not been examined. The other 59 children were selected for study without regard to the mental condition of the parents.

The *IQ* reported is the 1916 Stanford-Binet in every case but one, where the Kuhlman revision was used. Children 12 or over were generally examined before placement in a foster home but after dependency had been established. Children under three or four were generally not examined for one or two years. Almost all the other children were examined shortly after placement.

Skeels and Fillmore (4) have previously indicated a negative relation between the length of time a child spends in his own econom-

---

[2]These foster homes are for the most part boarding homes. There are a few free homes but no adoptive homes.

[3]A more detailed discussion of the mental development of this group will appear soon (6).

ically underprivileged home and the *IQ*. Our own data for the 184 children show an almost identical trend, although the decrease in *IQ* from children placed before three to children placed after 12 is somewhat sharper Table 1 shows a progressive decline in median

TABLE 1

Age at Placement and Binet *IQ*

Age at Placement	Children of feebleminded mothers *N*	Children of feebleminded mothers Md *IQ*	Children with two tests median *N*	Children with two tests median Md *IQ*	Children of normal mothers median *N*	Children of normal mothers median Md *IQ*	All children *N*	All children Md *IQ*
0- 2	12	100.5	6	102.5	15	99.9	27	100.3
3- 5	19	83.7	20	84.5	29	90.6	48	87.1
6- 8	12	74.6	18	88.3	36	90.1	48	82.3
9-11	9	71.5	8	88.7	22	91.8	31	81.6
12-15	16	53.1	7	79.2	14	80.7	30	66.9
Total	68		59		116		184	

IQ from 100.3 for children placed in foster homes before three to 66.9 for children placed after 12

Similar results were obtained when we considered the 68 children of mentally deficient mothers, the median *IQ* falls from 100.5 for children placed before three to 53.1 for those placed after 12

The same trend is evident in the median *IQ*'s of the 116 other children; there is a progressive decrease in median *IQ* from 99.9 for children placed before three to 80.7 for children placed after 12 Fifty-nine of these children have been reexamined since placement in the foster homes. There is no significant difference between these children and the group as a whole The median *IQ* of the children placed before three was 102.5, and there is a decrease to a median *IQ* of 79.2 for those placed after 12

The foster homes in which these children have been placed are definitely superior to their own homes and, in general, are average or superior homes in the community [4] They are, in the first place, economically stable; there is no mental deficiency or instability so far as we are able to discover, in most cases both parents are in

[4] This is the subjective estimate of the trained social workers on the staff of the Children's Service League based upon their familiarity with the homes in the community

the home, though occasionally a widow and a grown son are used; so far as investigation can determine, there is no immorality, alcoholism, or other social deficiency in the foster family. They are held in good repute by their neighbors. The children are given regular medical and dental attention by the agency, are well fed, clothed, and housed, and attend school regularly. If intelligence can be developed or improved by stimulating environmental situations, as Wellman suggests (7), we believe that it should occur in this situation, after placement in this superior environment. And, as measured by the Binet *IQ*, it does.

There is a median gain of 5.1 points for the group as a whole (see Table 2). Children placed after 12 made the greatest gain

TABLE 2
AGE AT PLACEMENT AND GAIN

Age at placement	Number	First median *IQ*	Second median *IQ*	Gain
0- 2	6	102.5	109.5	7.0
3- 5	20	84.5	89.6	5.1
6- 8	18	88.3	88.1	—.2
9-11	8	88.7	92.2	3.5
12-15	7	79.2	89.2	10.0
	59	88.6	93.7	5.1

(9.0 points). Children placed before three made the second largest gain (7.0 points). Children placed between three and five, and nine and eleven, made smaller gains (5.1 and 3.5 points respectively), but the children placed between six and eight showed a slight loss (.2 points).

We may consider the change in *IQ* in relation to the level on the first test, as shown in Table 3. Grouping the children on the basis of the first examination, we have six who are below 70; 29

TABLE 3
INITIAL LEVEL AND AMOUNT OF GAIN

Classification	Number	First *IQ*	Second *IQ*	Gain
Below 70	6	63.0	63.5	.5
70-89	29	80.2	87.7	7.5
90 and above	24	105.1	108.1	3.0

between 70 and 89, and 24 who are 90 or above On the second test, after placement in a foster home, those who were below 70 on the first test have made an average gain of 5 points, those who were between 70 and 89 have gained on the average, 7 5 points, and those who were 90 or above on the first test gained an average of 3 0 points

Table 4 presents the data somewhat differently. On the first test,

TABLE 4
CLASSIFICATION ON FIRST AND SECOND TEST

Classification on first test	Classification on second test: Defective	Dull	Average	Superior	Total
Defective	4	2			6
Dull	2	17	10		29
Average		4	14	5	23
Superior				1	1
Total	6	23	24	6	59

six children are classified as defective—below 70 *IQ* On retest, four of these are still defective, but two have gained enough to be classed as dull (9 and 11 points respectively) Twenty-nine children were classed as dull on the first test, but, on the second examination, two have lost enough (19 and 5) to be classed as defective, and 10 are now classed as average Of the 23 children who were first classed as average, four are now dull, 14 average, and five superior The only child who was classed as superior on the first examination is still superior and has gained 14 points in *IQ*

We have also considered the relation between birth order and the behavior of the *IQ* (Table 5) We classified our children into three

TABLE 5
BIRTH ORDER AND BINET *IQ*

Sibling order	Number	Median age	First median *IQ*	Second median *IQ*	Gain
Oldest	15	9 2	86 6	94 9	8 3
Middle	34	8 0	87 7	89 6	1 9
Youngest	10	7 4	91 5	96 5	5 1

groups, oldest and only child, middle, and youngest child. As we would expect from the previous analysis of the data, the youngest

child has the highest median *IQ* (91.5), the middle child is next (87.7), and oldest child is the lowest (86.6). These differences are not significant.

On the reexamination, the middle child made the smallest gain (1.9); the youngest child gained more than twice as much (5.1), and the oldest child gained more than four times as much (8.3) as the middle child.

We may summarize the results of this study briefly:

1. We have found a direct, negative relation between the length of time a child spends in an economically underprivileged home and the Binet *IQ*. When the home is particularly deficient, as in the case where the mother is mentally deficient, this relationship is even more marked.

2. When placed in foster homes superior to their own homes, the group, as a whole, has gained 5.1 points. Children placed after 1.2 and before three made the largest gains.

3. Children who were defective on the first examination tend to remain defective, the dull children made large gains and the average children gained consistently, but in a lesser degree.

4. When considered in relation to birth order, the oldest and youngest children made the greatest gains, the middle child gaining only slightly.

5. The data indicate that the *IQ* of the child is directly affected by environmental conditions, it is depressed by inferior or limited environmental stimulation, and increases when the child is placed in a superior environment.

There are several limitations to the present study, and we believe these should be further investigated.

1. The interval between tests should be important, if children gain as a result of placement in the foster home. Our cases are too few to permit thorough investigation of this aspect of the problem. For the present, we can only say that the test-retest interval for the group shows no relation to gain on retest ($r = .02$). We should like to know the effect of varied intervals between tests, for children placed in foster homes at the same age.

2. We should like to know the effect of varied intervals between tests for children of the same initial level of *IQ*.

3. The effect of different placement ages for children of the same initial *IQ* level should be investigated.

4 We need to study more carefully the level of the foster home, the discrepancy between this level and that of the true home, and the relationship between these factors and those already studied

5 We should know whether the sibling order in the true home is maintained in the foster home and, if it is not maintained, the relationship this has to mental development

REFERENCES

1. SIMPSON, B R The wandering *IQ* Is it time for it to settle down? *J of Psychol*, 1939, 7, 351-367
2. SKEELS, H M Mental development of children in foster homes. *J Genet. Psychol*, 1936, 49, 91-106
3. ———. Mental development of children in foster homes *J Consult Psychol*, 1938, 2, 33-43.
4. SKEELS, H M., & FILLMORE, E A The mental development of children from underprivileged homes *J Genet Psychol.*, 1937, 50, 427-439
5. SKODAK, M The mental development of adopted children whose true mothers are feebleminded *Child Devel*, 1938, 9, 303-308
6. SPEER, G S The mental development of children of feebleminded and normal mothers To appear in the *1940 Yearbook, Nat Soc Stud Educ*
7. WELLMAN, B L. Mental growth from preschool to college *J Exper Educ*, 1937, 6, 127-138
8. ——— Guiding mental development *Childhood Educ*, Nov, 1938, 108-112

Child Guidance Service of Sangamon County
717 South Grand Avenue, East
Springfield, Illinois

The Journal of Genetic Psychology, 1940, 57, 57-63

STUDIES IN ANIMISM III ANIMISM IN FEEBLE-MINDED SUBJECTS*

Departments of Psychology, University of Nebraska and University of Virginia, and the Iowa Child Welfare Research Station

ROGER W RUSSELL, WAYNE DENNIS, AND F ELTON ASH

A INTRODUCTION

Russell and Dennis (3) have described a standardized procedure which makes it possible to classify children into the four stages of animism outlined by Piaget (1) and, by so doing, permits an objective study of the development of animism.

The senior author (2) has investigated the developmental course of animism in a group of 774 normal children and obtained results that in general corroborate Piaget's analysis These subjects were divisible into three groups—urban, suburban and rural—so selected because of the belief that the development of animism might be influenced by experience with or acquaintance with natural phenomena. In this regard one would expect the suburban children to have more contact with nature than the urban children and the rural children more than either of the other two groups. Comparable mental and chronological age levels were studied and the fact that there were no significant differences in the development of animism between the three groups and that where the differences did approach significance they favored the urban group over the suburban and rural groups indicate that this particular type of experience is not an influencing factor

The present experiment is concerned in a still different manner with the study of the effects of experience on the development of animism Feeble-minded subjects have been selected for examination, since, by comparing these subjects with normal children of the same mental age, it will be possible to study the influence of additional years of experience on the manifestations of animism when mental age is held constant

*Received in the Editorial Office on May 27, 1939

B. Subjects[1]

Four hundred and thirty feeble-minded subjects were examined during the progress of the experiment. Two hundred and fifty-five of these were patients at the Lynchburg State Colony in Virginia, and the other 175 resided at the Wrentham State School for the Feeble-Minded in Massachusetts. The Virginia group, which we shall refer to as the "*L*" group, ranged in mental age from 2 years 11 months to 11 years 7 months and in chronological age from 8 years 0 months to 64 years 10 months. The Massachusetts group, hereafter spoken of as the "*W*" group, ranged in approximately equal numbers at each age level from a mental age of four through a mental age of ten and were all within a chronological age range of from 7 years 2 months to 20 years 10 months.

Some of the "*L*" group subjects and most of the "*W*" group attended special classes and secured as much educational training as their ability warranted, and all but the most severely handicapped had occupational duties to perform. The amount of time that the subjects had lived at the institutions varied considerably, some having spent a great amount of their lives there and others having been recently admitted.

The subjects had all been examined by the staff psychologists soon after admission to the institutions. Since these tests, individual Stanford-Binets, had been given over a period of several years it was necessary to obtain the mental age of each subject at the time of our examination. This was accomplished by converting the usual form of the equation for deriving the Intelligence Quotient to the form

$$MA = IQ \times CA$$

Thus, knowing the Intelligence Quotient of the subject at the time of intelligence testing and the chronological age of the subject at the time of examination for stage of animistic concept, it was possible to obtain his mental age at the time of our questioning. This pro-

[1]The writers wish to acknowledge the assistance rendered by Dr. G. B. Arnold, Superintendent of the Lynchburg State Colony for Epileptics and Feeble-Minded, Dr. W. S. Raymond, Superintendent of the Wrentham State School for the Feeble-Minded, and the other members of the staff at these two institutions for their valuable assistance during the collection of the data. Particular thanks are due Mr. John N. Buck and Miss Ruth Prouty, psychologists at the two institutions respectively, for the use of their intelligence test results.

cedure is based upon two assumptions, for which there is considerable evidence. First, it is necessary to assume that the Intelligence Quotient is relatively constant in the feeble-minded; and, second, a basal chronological age of 14 is taken for all subjects 14 or above.[2]

The "*L*" group subjects were divided into two groups on the basis of chronological age, the "*L*-1" group (130 subjects) being below 21 years of age and Group "*L*-2" (125 subjects) being 21 or above. This was done in order to compare the "*L*-1" group with the "*W*" group, all of whom were below 21 years in chronological age.

The method described by Russell and Dennis (3) and employed in the study of normal children was used throughout this research.

C. Results

1. *Classification into Stages*

Of the 430 feeble-minded subjects examined only five had systematic distinctions between the animate and the inanimate which were not in accord with any of the four stages. This means that of the subjects other than those in the No Concept group, of which there were 98, 98.5 per cent of the feeble-minded were classifiable into the designated stages of animism.

2. *Correlation of Stage with Mental Age*

The coefficients of Mean Square Contingency for the relationships between mental age and the stage for the three feeble-minded groups and the combined normal group are given in Table 1.

TABLE 1
Coefficients of Mean Square Contingency for the Relationship between *MA* and Stage

Group	"*C*"
"*W*"	0.71
"*L*-1"	0.62
"*L*-2"	0.61
Combined normal	0.59

The fact that the coefficient which was obtained for the "*W*" group is higher than that which characterized the other groups is probably due to the fact that there were more cases at the lower mental age levels in Group "*W*." This group was selected so as to

[2]Subjects who had been tested at ages beyond 14 years required no treatment of their *M.A.* scores. These comprised a large number of the subjects.

contain approximately 25 cases at each mental age level from four to ten, whereas the other groups contained relatively few cases below the six year mental age level.

Considering the fact that the animism examination is a relatively short procedure, often taking no more than five minutes, the correlations between Stage and mental age are surprisingly high. Nevertheless, they are not high enough to permit accurate individual prediction of Stage from mental age or of mental age from Stage.

3. *Comparison of Feeble-Minded and Normal Subjects*

For comparison with the normal children there are sufficient subjects to warrant treatment in terms of percentages only for mental ages 6^0 to 9^{11} inclusive. This interval includes 66 per cent of all the cases.

Figure 1 shows the progression of stage with mental age for the

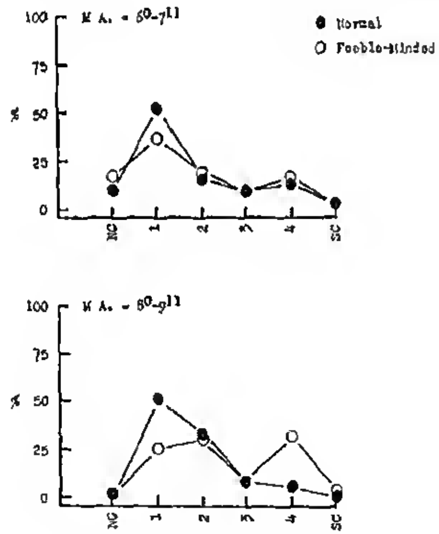

FIGURE 1

6^0-7^{11} and 8^0-9^{11} mental age intervals. At the younger mental age level the feeble-minded and normal groups are essentially similar. There are more feeble-minded at the No Concept level, the Critical Ratio for the difference between the groups being 2.01, but, con-

versely, there are more normals at the Stage 1 level, the Critical Ratio of the difference being 2.50. At the Stage 2, 3, and 4 levels and in the Special Concept group both feeble-minded and normal subjects are equally distributed.

Significant differences appear at the 8^0-9^{11} mental age interval and favor the feeble-minded group as more advanced in animistic concept. At this interval there are more normal than feeble-minded subjects at the Stage 1 level and more feeble-minded than normals at the Stage 4 level, the Critical Ratios of these differences being 4.26 and 5.57 respectively.

4. *Effects of Advanced Chronological Age*

In order to determine the effects of the added years of experience of the feeble-minded subjects, comparisons were made at mental age levels of 6^0-7^{11} and 8^0-9^{11} between the feeble-minded subjects under 21 years chronological age and those 21 or above.

Figure 2 shows the progression of stages for these groups. It is immediately apparent that the subjects 21 years or above in chronological age are more advanced than those under 21. The Critical

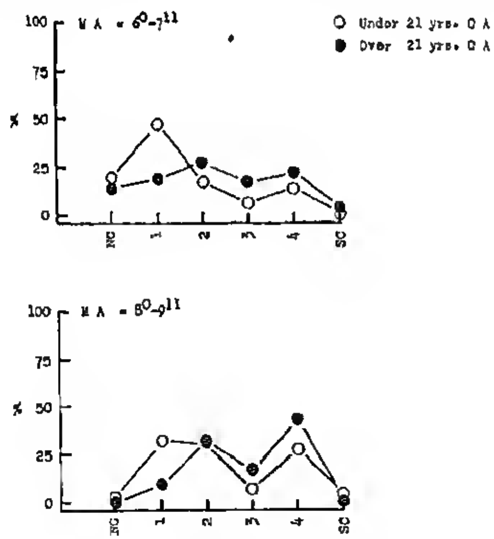

FIGURE 2

Ratios for the differences at the 6^{0}-7^{11} mental age interval at the Stage 1, Stage 2, Stage 3, and Stage 4 levels are 3.95, 1.51, 1.69, and 1.33 respectively and at the 8^{0}-9^{11} mental age interval at the Stage 1, Stage 3, and Stage 4 levels, 3.36, 1.62, and 1.68 respectively.

The differences between the two groups of subjects under 21 years of chronological age, the "*L-1*" and the "*H*" groups, were in no instance reliable.

D. Summary and Discussion

The data here presented show clearly that subjects who have been at a given mental age for several years are more advanced in regard to their concepts of animation than children who have just reached the same mental level. It has also been shown that the older feeble-minded subjects, who have been at a given mental age for a considerable period of time, are more advanced in the animistic scale than are the younger feeble-minded subjects. In other words, these approaches have indicated that with subjects of the mental levels here dealt with, age is a variable which affects the development of ideas when mental age is held constant by an appropriate selection of cases.

It may be suggested as an alternative explanation of the advancement of the feeble-minded subjects that they live in a rural environment. The Wrentham and the Lynchburg colonies are farm colonies, and nearly all members participate in farm activities. In this connection, attention must be called to the fact that one of our normal groups also was decidedly rural in character. Since no differences were found between normal rural and urban groups the differences between the rural normal group and the feeble-minded subjects is as great as the differences between the combined normal groups and the feeble-minded. This consideration shows that an explanation of terms of rural environment is not tenable.

We would suggest that added years of experience affect the concepts of the institutional subject by providing a greater likelihood that the subject will notice characteristics of objects which have previously escaped his attention, and by providing increased opportunities to hear and to assimilate adult concepts. This statement, however, must be taken as a suggestion and not as a conclusion. In no sense can it be said that we have observed the processes by which an older subject acquires more mature ideas.

Thus far we have dwelt upon the positive rôle played by age or experience in the development of animistic concepts, but our data also show the great limitation which mental age places upon the effectiveness of experience. Our figures have compared the feeble-minded subjects with children of the same mental age. If, on the other hand, we compare feeble-minded adults with normal adults we see how relatively ineffective is experience when the mental level is low.

Studies in progress show that practically all adults of average intelligence are in Stage 4. In contrast to this, the adult feeble-minded subjects of mental age 6^{0}-7^{11} are in Stage 4 in only 22 per cent of the cases, and those of the 8^{0}-9^{11} level are in Stage 4 in only 43 per cent of the cases. The comparable figures for normal children are 13 per cent and 6 per cent (2). While the comparison between the feeble-minded adults and the normal children of the corresponding mental age shows the feeble-minded to be more advanced than the children, it must be borne in mind that in regard to their ideas of animation the feeble-minded adults are nearer to the children than they are to the adults.

References

1. Piaget, J. The Child's Conception of the World. New York: Harcourt, Brace, 1929. Pp. ix + 397.
2. Russell, R. W. Studies in animism: II. The development of animism. *J. Genet. Psychol.*, 1940, 56, 353-366.
3. Russell, R. W., & Dennis, W. Studies in animism: I. A standardized procedure for the investigation of animism. *J. Genet. Psychol.*, 1939, 55, 389-400.

Department of Psychology
University of Nebraska
Lincoln, Nebraska

Department of Psychology
University of Virginia
University, Virginia

Iowa Child Welfare Research Station
Iowa State University
Iowa City, Iowa

The Journal of Genetic Psychology, 1940, 57, 65-82

THE RELATIONSHIP OF *IQ* TO HEIGHT AND WEIGHT FROM THREE TO FIVE YEARS*

Developmental Health Inquiry, Laboratory of Anatomy, Western Reserve University

EVELYN KATZ[1]

A. THE PROBLEM

The present investigation is an attempt to determine the relationship at the preschool level of height and weight to intelligence as measured by the Stanford-Binet test. Specifically, the problem resolves itself into two questions: (*a*) What is the correlation of height and weight with *IQ* during the period from three to five years? (*b*) If children of these ages are divided into "very superior," "superior," and "average" groups on the basis of *IQ*, will the groups show differences in height and weight?

Beginning with Porter's investigation in 1892, numerous attempts have been made to estimate the degree of relationship that exists between mental and physical traits. The majority of these studies can be criticized either because of inadequate statistical treatment, lack of a satisfactory criterion of intelligence, or failure to control such factors as age, sex, race, and socio-economic status. A comprehensive and critical review of the literature in this field up to 1930 is given in Paterson's *Physique and Intellect* (7). Paterson concludes that the general trend is toward a small positive correlation between mental and physical traits. Jones, in two subsequent reviews (5, 6), surveys the literature up to 1936 and arrives at substantially the same conclusion. Paterson noted a dearth of evidence on the relationship between physical and mental traits below school age. Moreover, the situation has not materially changed

*Recommended for publication by Dr. Theodore T. Zuck, and received in the Editorial Office on June 6, 1939.

[1]The writer wishes to acknowledge her indebtedness to Dr. L. Dewey Anderson for many valuable suggestions during the course of the study, to the late Dr. T. Wingate Todd for making the physical records available, and to the staff members responsible for collecting the physical and psychological data.

since 1930 Only two studies which attack the problem at the preschool level have come to the writer's attention

Bayley (1) reports a study of approximately 50 children who were observed at regular intervals up to six years of age None of the physical measurements showed any significant relationship to the psychological measurements

Honzik and Jones (4) obtained mental and anthropometric measurements on 127 boys and 125 girls between the ages of 21 months and 7 years, comprising a representative sampling of the population of Berkeley, California They used the California Preschool Mental Scale through five years and the Stanford-Binet at six and seven years The correlations of height and mental test scores ranged from .07 at three years to 21 at seven years for girls, and from .02 at three years to 18 at seven years for boys For weight and mental test scores, the correlations over this same period ranged from — 04 to .12 for girls, and from — 09 to 20 for boys.

It is reasonable to assume that the wealth of material on the older as compared with the younger ages is not an indication of the relative significance of the problem at the two periods The period below six years would appear to be an important one to investigate whenever the problem is that of determining relationships between different phases of development. It is possible that during these years, when growth is progressing at a rapid rate, the relationship between traits is different from that found during periods of slower growth. There is certainly no justification for assuming that the slight correlations between mental and physical traits found above six years would imply a similar relationship below six. One may conclude, then, that the small amount of material on the preschool years is not due to the unimportance of the problem but rather to the inaccessibility of subjects The longitudinal records of physical and mental growth collected by the Developmental Health Inquiry make possible contributions to this field of investigation It is hoped that the study may be continued as data are collected on these children at later ages

B The Method

The children used in this study are 112 boys and 117 girls from the longitudinal growth program of the Developmental Health In-

quiry Enrollment on the series is accomplished through the pediatrician, who recommends a child for observation only if he is free from gross physical and mental defect As one might expect where voluntary cooperation of parent and pediatrician is the basis for selection, the subjects are drawn primarily from the upper socio-economic groups in the community In racial background the children are, with few exceptions, of American-born parents of North European stock

The program of the Inquiry involves periodic measurement of the child at three-month intervals during the first year, semi-annually from one to five years, and annually after five years Each child is examined as near as possible to his birthdate and to the exact three- or six-month interval between. In practically all instances examinations are made within a period of two weeks before or after the exact birthdate For the present investigation, children were selected who had had regular examinations at six-month intervals from three to five years inclusive In a few instances, however, one or more complete examinations were missed

This study was limited to the three- to five-year-old children for two main reasons First, there was a large group of children who had received regular physical and mental examinations at these ages Secondly, the old form of the Stanford-Binet test had been used consistently during this period The study was cut off at three years since below that age a Binet *IQ* is not a satisfactory measure of mental status, it was impossible to continue the study beyond five years because a large proportion of the children had not yet reached their sixth birthday

Among other measures the following data were available for each child at each age studied Stanford-Binet *IQ*, height in millimeters, weight in pounds

The psychological examinations were administered by experienced examiners, all of whom had had graduate work in psychology as well as testing experience. The physical measurements were made either by a physician or under the supervision of a physician on the Inquiry staff

Two general techniques were used in analyzing the data. the correlational technique and the group-difference approach The first method involved obtaining correlations between height and *IQ* and

weight and *IQ* for each sex at each age Since this is the usual method of studying relationship, it permits comparisons between the results of the present study and those of other investigators. However, it has the disadvantage of requiring the use of every psychological test result obtained on the child It is not uncommon to find, especially in very young children, one or two test scores which can in no sense be regarded as representative of the child's ability When such scores are included in correlations, the unreliability of the mental measurements may affect the size of the correlations It seems reasonable to expect a more reliable measure of a child's ability when several examinations are considered than when an estimate is based upon the results of a single examination In order to test the hypothesis that the correlations of *IQ* with height and weight might be raised by adopting a more reliable measure of intelligence, they were computed by using for each child the median of five *IQ*'s obtained in examinations from three to five years. The correlations secured by this procedure were then compared with those using the *IQ* at each age as the variable.

A second attack on the problem was made by the method of group differences, on the assumption that even when the correlation between two variables is small, groups that are at opposite extremes in one variable may show a corresponding difference in the other variable That is, even if a relatively low correlation were found between height and intelligence, a comparison of the height of two groups widely separated in degree of ability might yield a significant difference Consequently, for the present study, three groups of children, an "average," a "superior," and a "very superior" group were used In order to make fairly certain that the groups really represented different degrees of mental ability, the median *IQ* of five examinations rather than the results of a single examination was again adopted as the better measure

A total of 135 subjects was selected, divided into three groups of 45 each, on the basis of ability (*a*) a "very superior" group composed of 21 boys and 24 girls whose median *IQ*'s on examinations from three to five years were over 125, (*b*) a "superior" group, composed of 20 boys and 25 girls whose median *IQ*'s on examinations from three to five years were between 110 and 119, (*c*) an "average" group composed of 25 boys and 20 girls whose median *IQ*'s on examinations from three to five years were between 90 and 105

It was impossible to secure enough cases from the population studied for an adequate sampling of a below-average group

C. The Results

1. *The Correlational Method*

The correlations of *IQ* with height and weight are given in Tables

TABLE 1

CORRELATIONS BETWEEN HEIGHT AND ACTUAL *IQ* AND HEIGHT AND MEDIAN *IQ* FROM THREE TO FIVE YEARS

Age	Male *N*	Male *R*	Male *PE*	Female *N*	Female *R*	Female *PE*
			Actual IQ			
3 yrs	100	—083	±067	107	312	±059
3½ yrs	108	—050	±065	114	268	±059
4 yrs.	107	—042	±064	114	278	±058
4½ yrs	108	004	±065	115	362	±055
5 yrs	105	027	±066	110	220	±061
			Median IQ			
3 yrs	107	—069	±065	115	399	±053
3½ yrs.	108	—071	±065	115	366	±.055
4 yrs	107	—029	±065	114	359	±055
4½ yrs	108	—019	±065	115	385	±054
5 yrs	105	019	±066	114	394	±053

TABLE 2

CORRELATIONS BETWEEN WEIGHT AND ACTUAL *IQ* AND WEIGHT AND MEDIAN *IQ* FROM THREE TO FIVE YEARS

Age	Male *N*	Male *R*	Male *PE*	Female *N*	Female *R*	Female *PE*
			Actual IQ			
3 yrs	104	011	±066	110	180	±062
3½ yrs	111	—075	±064	115	284	±058
4 yrs	112	—070	±063	116	288	±057
4½ yrs	112	006	±064	117	222	±059
5 yrs	106	—017	±065	111	257	±.060
			Median IQ			
3 yrs	111	—042	±064	115	310	±057
3½ yrs	111	—063	±064	117	328	±056
4 yrs	111	—054	±064	115	344	±056
4½ yrs	112	001	±061	115	315	±057
5 yrs	107	—025	±065	113	266	±059

1 and 2, using as the measure of intelligence both the actual *IQ* obtained at each age and the median *IQ* of the examinations from three to five years. Table 3 is an evaluation of the differences between the correlations of the two sexes.

TABLE 3

DIFFERENCES BETWEEN BOYS AND GIRLS IN THE CORRELATIONS OF *IQ* WITH HEIGHT AND *IQ* WITH WEIGHT

	Actual *IQ*			Median *IQ*		
Age	*Diff*	PE_{diff}	$\frac{Diff}{PE_{diff}}$	*Diff*	PE_{diff}	$\frac{Diff}{PE_{diff}}$
			Height			
3 yrs.	.393	±.089	4.42	.468	±.084	5.57
3½ yrs.	.318	±.087	3.66	.437	±.085	5.14
4 yrs.	.320	±.087	3.68	.388	±.085	4.56
4½ yrs.	.358	±.085	4.21	.404	±.084	4.81
5 yrs.	.193	±.090	2.14	.375	±.085	4.41
			Weight			
3 yrs.	.139	±.091	1.53	.352	±.086	4.09
3½ yrs.	.359	±.086	4.17	.391	±.085	4.60
4 yrs.	.358	±.085	4.21	.398	±.085	4.68
4½ yrs.	.216	±.087	2.48	.304	±.086	3.53
5 yrs.	.274	±.088	3.11	.291	±.088	3.31

The most striking result found in these tables is the marked sex difference in the magnitude of the correlations. For the girls the correlations between the physical and psychological measures are all low but positive. The corresponding correlations for the boys hover around zero, with a larger number of negative than of positive correlations. The critical ratios of these sex differences are uniformly high when the median *IQ* is employed, indicating, at each age for both height and weight, 99 or more chances in 100 that there is a true sex difference in the degree of relationship of the physical to the psychological measures. The sex differences are less pronounced when the actual *IQ* is used as a variable, the critical ratios in several instances falling well below the criterion of statistical significance. Nevertheless, even with this measure of mental ability there is no single case in which the correlation for the boys equals or exceeds that for the girls.

The sex difference may also be stated in another way. The probable error of a correlation of zero for 113 to 117 cases, the number of girls on which the correlations of median *IQ* with the physical

measures were based, is between .062 and .063. At each age studied the correlations of height and weight with median *IQ* are greater than four times this probable error, which means that there are more than 99 chances in 100 that the true correlations between these measures are greater than zero. For the boys, on the other hand, the correlations are all close to zero or have slightly negative values, showing a complete lack of relationship between the variables.

The same trend is present in the correlations of height and weight with the actual *IQ* at each age. With the exception of the three-year correlation between weight and *IQ* and that at five years between height and *IQ*, the correlations for the girls are all greater than four times the probable error of zero correlation. For the boys the evidence is consistently in the direction of a lack of correlation between *IQ* and the physical measures.

Comparing the correlations obtained when the median *IQ* is used

TABLE 4

DIFFERENCES BETWEEN CORRELATIONS OF ACTUAL *IQ* WITH HEIGHT AND WEIGHT AND OF MEDIAN *IQ* WITH HEIGHT AND WEIGHT*

Age	Male Diff**	PE_{diff}	$\frac{Diff}{PE_{diff}}$	Female Diff**	PE_{diff}	$\frac{Diff}{PE_{diff}}$
			Height			
3 yrs	.014	±.034	.41	.087	±.035	2.49
3½ yrs	−.021	±.032	−.66	.098	±.033	2.97
4 yrs	.013	±.029	.45	.081	±.029	2.79
4½ yrs	−.023	±.027	−.85	.023	±.029	.79
5 yrs	−.008	±.034	−.24	.174	±.034	5.12
			Weight			
3 yrs	−.083	±.033	−2.52	.130	±.037	3.51
3½ yrs	.012	±.036	.33	.044	±.036	1.22
4 yrs	.016	±.029	.55	.056	±.029	1.93
4½ yrs	−.005	±.025	−.20	.093	±.027	3.44
5 yrs	−.008	±.035	−.23	.009	±.036	.25

*The formula used for calculating the probable error of the difference was $PE_{diff} = PE_1^2 + PE_2^2 - 2rPE_1PE_2$ where

PE_1 = the probable error of the correlation using the actual *IQ*

PE_2 = the probable error of the correlation using the median *IQ*

r = the correlation between the actual *IQ* and the median *IQ*

** + indicates that the correlation is higher with the median *IQ*

— indicates that the correlation is higher with the actual *IQ*

with those obtained with the actual IQ, it is apparent that for girls the former measure yields higher correlations at every age for both height and weight. In Table 4 these differences are evaluated. In spite of the many low critical ratios, it is worth noting that no exception occurs in the direction of the difference. The data for the boys, on the contrary, present no consistent trend, many of the correlations being higher with the actual IQ than with the median IQ. Thus it appears that where some relationship exists between the IQ and another variable, as it does for the girls with both height and weight, using a more reliable measure of IQ than that obtained on a single examination tends to raise the correlation.

2. *Method of Group Differences*

The following tables summarize the results of a comparison of the physical measurements of the three ability groups. Table 5 presents the means and standard deviations in height (expressed in terms of standard scores) for the ability and sex groups at each age. Raw scores were converted into standard scores because it was the writer's original intention to combine the data of the two sexes. When marked sex differences were noted, this became impossible. The formula used to obtain the standard score was $\frac{X-M}{SD}$ in which X refers to the raw score, and M and SD refer to the mean and standard deviation of the particular age-sex group based upon all the Inquiry cases. To facilitate calculations each sigma score was increased by five. In Table 6 the differences between the ability groups and the significance of these differences are given.

It is again evident that the relationship between height and mental test score differs for the two sexes. For boys the critical ratios are so small as to indicate no real differences between the ability groups. In fact, in the comparison of the superior with the average boys, the small differences obtained are in favor of the lower intelligence group. In striking contrast to these negative results is the picture of clean-cut differences for the girls. A comparison of the "very superior" with the "average" group yields critical ratios of over three at each age, and considering the small number of cases, the ratios obtained in the comparison of the "superior" and "average" groups are of sufficient magnitude to show a definite trend. The smaller ratios for the differences between the "very superior" and "superior" groups point merely to a tendency for the results

TABLE 5

THE MEANS AND STANDARD DEVIATIONS IN HEIGHT OF EACH ABILITY AND SEX GROUP AT EACH AGE

Age in years	"Very Superior" N	Mean	SD	"Superior" N	Mean	SD	"Average" N	Mean	SD
				Male					
3	21	5 1726	1 0483	20	4 9625	8414	25	5 0350	9769
3½	21	5 1490	1 0145	20	4 9250	9274	25	5 0950	9548
4	21	5 0500	1 0493	20	4 8625	9303	25	5 0450	9320
4½	21	5 0890	.9521	20	4 8250	8718	25	5 0150	8862
5	21	5 0890	.9521	19	4 8880	9983	23	5 0050	8561
				Female					
3	24	5 4792	9866	25	5 1150	8674	20	4.2560	1 1589
3½	24	5 5310	1 0073	25	5 1450	9244	19	4 3090	1 1833
4	23	5 4620	9314	25	5 1850	9494	20	4 3375	1 0615
4½	23	5 4400	9869	25	5 0950	.8981	20	4 3375	1 0139
5	23	5 5050	.9942	25	5 1550	9309	20	4 3625	1 0532

TABLE 6

DIFFERENCES IN HEIGHT BETWEEN "VERY SUPERIOR," "SUPERIOR," AND "AVERAGE" GROUPS

Age	Diff*	SD_{diff}	$\frac{Diff}{SD_{diff}}$
		Male	
		"Very superior" and "Superior"	
3 yrs	+ 2101	2960	71
3½ yrs	+ 2240	3034	74
4 yrs	+ 1875	3135	60
4½ yrs	+.2610	2819	93
5 yrs	+ 2010	3092	65
		"Very superior" and "Average"	
3 yrs	+ 1376	3007	46
3½ yrs	+ 0540	2924	18
4 yrs	+ 0050	2996	02
4½ yrs	+ 0740	.2731	27
5 yrs.	+ 0840	3065	28
		"Superior" and "Average"	
3 yrs	— 0725	2711	.27
3½ yrs	— 1700	2819	60
4 yrs.	— 1825	.2793	.65
4½ yrs.	— 1900	2634	72
5 yrs	— 1170	3213	.36
		Female	
		"Very superior" and "Superior"	
3 yrs	+ 3642	2660	1 37
3½ yrs.	+ 3860	2765	1 40
4 yrs	+ 2770	.2716	1 02
4½ yrs.	+.3450	2731	1 26
5 yrs	+.3500	2786	1 26
		"Very superior" and "Average"	
3 yrs	+1 2232	3232	3.85
3½ yrs	+1 2220	3406	3 59
4 yrs	+1.1245	3067	3 67
4½ yrs	+1 0525	3062	3 44
5 yrs.	+1 1425	3137	3 64
		"Superior" and "Average"	
3 yrs	+ 8590	3066	2.77
3½ yrs.	+ 8360	3285	2.54
4 yrs	+ 8475	3040	2 79
4½ yrs	+ 7075	2892	2 45
5 yrs	+.7925	.3002	2.64

* + indicates that the difference is in favor of the higher intelligence group.
— indicates that the difference is in favor of the lower intelligence group.

TABLE 7

THE MEANS AND STANDARD DEVIATIONS IN WEIGHT OF EACH ABILITY AND SEX GROUP AT EACH AGE

Age in years	"Very Superior" N	"Very Superior" Mean	"Very Superior" SD	"Superior" N	"Superior" Mean	"Superior" SD	"Average" N	"Average" Mean	"Average" SD
					Male				
3	21	4 9703	1 0220	20	5 0500	9223	25	4 9350	.9281
3½	21	5 0655	9126	20	5 0375	9229	25	4 8750	8426
4	21	4 8625	9942	20	5 0125	9907	25	4 8950	7098
4½	21	4 9822	9685	20	5 1750	1 1000	25	4 8850	8986
5	21	4 8988	9060	19	5 1250	9389	23	4 8968	7558
					Female				
3	24	5 2708	1 0990	25	4 9450	1 3537	20	4 2625	1 1053
3½	20	5 3289	1 0103	25	5 0550	1 0162	19	4 2829	6940
4	23	5 2554	9502	25	5 0650	9858	20	4 3250	7580
4½	23	5.1902	1 0033	25	4 9650	1 0220	20	4 2750	6955
5	23	5 1359	8892	25	4 9450	9368	20	4 4750	7442

to be in the expected direction but they are too low to be of statistical significance.

TABLE 8

DIFFERENCES IN WEIGHT BETWEEN THE "VERY SUPERIOR," "SUPERIOR," AND "AVERAGE" GROUPS

Age	Diff*	SD_{diff}	$\frac{Diff}{SD_{diff}}$
		Male	
		"Very superior" and "Superior"	
3 yrs.	—.0797	.3037	.26
3½ yrs	+.0280	.2868	.10
4 yrs	—.1500	.3062	.49
4½ yrs	—.1928	.3243	.59
5 yrs	—.2262	.2924	.77
		"Very superior" and "Average"	
3 yrs	+.0353	.2901	.12
3½ yrs.	+.1905	.2608	.73
4 yrs	—.0325	.2610	.12
4½ yrs	+.0972	.2774	.35
5 yrs	+.0020	.2513	.01
		"Superior" and "Average"	
3 yrs	+.1150	.2774	.41
3½ yrs	+.1625	.2664	.61
4 yrs.	+.1175	.2675	.44
4½ yrs	+.2900	.3016	.95
5 yrs.	+.2282	.2655	.86
		Female	
		"Very superior" and "Superior"	
3 yrs	+.3258	.3514	.93
3½ yrs	+.2679	.2895	.93
4 yrs	+.1904	.2867	.66
4½ yrs	+.2252	.2925	.77
5 yrs	+.1909	.2636	.72
		"Very superior" and "Average"	
3 yrs	+1.0083	.3338	3.02
3½ yrs	+1.0400	.2605	3.99
4 yrs	+.9304	.2605	3.57
4½ yrs	+.9152	.2607	3.51
5 yrs	+.6609	.2509	2.63
		"Superior" and "Average"	
3 yrs	+.6825	.3665	1.86
3½ yrs.	+.7721	.2581	2.99
4 yrs.	+.7400	.2538	2.92
4½ yrs	+.6900	.2569	2.69
5 yrs	+.4700	.2523	1.86

* + indicates that the difference is in favor of the higher intelligence group
— indicates that the difference is in favor of the lower intelligence group

Table 7 summarizes the means and standard deviations in weight for the ability and sex groups at the various ages. In Table 8 the differences between the ability groups are evaluated. The results are essentially like those for height. Again there is no tendency for bright boys to be more developed physically than average boys, but bright girls tend to be heavier than average girls. Although the critical ratios for the weight differences of the female groups are less consistent from age to age and somewhat smaller on the whole than the height differences, the evidence is unmistakably in the expected direction.

The greatest difference between the ability groups in both height and weight is, of course, the difference between the "very superior" and "average" girls. These differences in terms of inches and pounds are indicated in Table 8a.

TABLE 8A

	Height differences	Weight differences
3 years	1.65 inches	3.14 pounds
3½ years	1.71 inches	4.51 pounds
4 years	1.63 inches	4.31 pounds
4½ years	1.63 inches	4.63 pounds
5 years	1.85 inches	4.16 pounds

It must be noted that the consistency of the results from age to age, both in the correlations and in the group differences, would be expected in view of the fact that the same children were followed from three to five years. Simmons and Todd (8) have found at these ages correlations for both height and weight of over .90 between measurements at successive examinations. Thus relationships found at three years would be reflected again at succeeding ages.

It is difficult to account for the marked sex difference appearing both in the correlation coefficients and in the group differences. The same method of selection was used for both sexes. The same criteria of physical and psychological health were used for both boys and girls on the longitudinal study. Under these conditions, the chance of differential selection seems extremely remote.

The possibility that the significant differences found for the girls between the three ability groups might be explained on the basis of socio-economic selection was explored. As has been pointed out

above, the sample of children studied by the Developmental Health Inquiry is relatively homogeneous with respect to socio-economic status since practically all the subjects come from the three upper socio-economic groups as defined by the Minnesota Occupational Classification (3, Appendix A). Nevertheless, if it were found that the majority of the "very superior" girls were in Group I while the "average" girls were predominantly from Group III, then the differences in the physical measurements might merely reflect socio-economic differences. However, the socio-economic distribution for the various sub-groups makes this explanation untenable (Table 8b)

TABLE 8B
Socio-Economic Group

	I		II		III	
	Boys	Girls	Boys	Girls	Boys	Girls
"Very superior"	52%	33%	29%	29%	19%	38%
"Superior"	55%	32%	20%	36%	25%	32%
"Average"	40%	45%	32%	25%	28%	30%

Group I represents children whose fathers are in the professions, Group II children of business men, and Group III children of clerical workers and skilled tradesmen

The above data show that there is in Group I a larger proportion of "very superior" boys than of "very superior" girls, whereas the reverse is true for Group III. For the "average" group, the proportion of girls in Group I is but slightly larger than the proportion of boys, with practically the same percentage of each sex in Group III. In the light of the above distribution, if the differences were due to socio-economic factors one might expect them to appear in boys rather than in girls.

Nor can the sex difference be explained on the basis of the relative maturity of the two sexes. It is true that girls have been found to mature more rapidly than boys For example at four years, girls, on the average, have attained approximately 27 per cent of their adult weight and approximately 61 per cent of their adult height, proportions not reached by the boys until five years (2, pp 210, 228-229). Nevertheless, if we compare the relationship of intelligence to height and weight for the two sexes when they are at the same level of maturity, we find that there does appear to be a cor-

TABLE 9

MEANS, STANDARD DEVIATIONS AND COEFFICIENTS OF VARIABILITY OF THE MALE AND FEMALE GROUPS FOR MEDIAN *IQ*, HEIGHT AT THREE YEARS AND WEIGHT AT FOUR YEARS

	Male				Female			
	N	Mean	*SD*	Coefficient Variability	*N*	Mean	*SD*	Coefficient Variability
Median *IQ*	112	109 11	14 46	13 25	117	111 73	12 59	11.26
Height at 3 years in mm	107	950 83	34 88	3 67	115	947 04	34 26	3 62
Weight at 4 years in lbs	112	36 44	4.79	13 14	115	35 87	4 66	13 00

relation for the girls at four years but no relationship for the boys at five years

Finally we must examine the hypothesis that the sex difference may be explained by a greater variability of the girls. The largest sex difference in the correlations was found at three years for height and at four years for weight. Table 9 presents the means, standard deviations, and coefficients of variability of the two sexes in median *IQ*, height at three years, and weight at four years. The coefficients of variability are shown to be larger for the boys. These data then obviously invalidate the explanation that the sex difference may be accounted for on the basis of a greater variability of the girls.

D. Summary and Conclusions

1. The present study was an attempt to determine the relationship of Stanford-Binet *IQ* to height and weight in children from three to five years of age

2. The children, totaling 112 boys and 117 girls who came predominantly from the three upper socio-economic groups as defined by the Minnesota Occupational Classification, were examined at six-month intervals during this age period.

3. Correlations between height and *IQ* and weight and *IQ* were calculated in two ways (*a*) by using the actual *IQ* at each age as the measure of intelligence, and (*b*) by using the median *IQ* of the examinations from three to five years as the variable. Comparisons were also made between three groups of children selected on the basis of the median *IQ*: a "very superior" group composed of children whose median *IQ*'s on examinations from three to five years were 125 or over, a "superior" group composed of children with median *IQ*'s between 110 and 119, and an "average" group composed of children with median *IQ*'s between 90 and 105.

4. For girls significant correlations were found at each age between *IQ* and height and *IQ* and weight both with the actual *IQ* and with the median *IQ* as the variable. On the other hand for the boys correlations between these same measures were all approximately zero

5. The sex difference in the size of the correlations between the physical and psychological measures was more pronounced when the median rather than the actual *IQ* was used

6. For the girls the correlations between the physical and psy-

chological measures were without exception higher when the median IQ than when the actual IQ was used The correlation coefficients for the boys, however, were in some instances higher with the median IQ and in others, higher with the actual IQ.

7 When groups of children of "very superior," "superior," and "average" intelligence were compared there was a tendency for the "very superior" girls to be taller and heavier than the "superior," and for the "superior" to exceed the "average" in these measurements The critical ratios for the differences between the "very superior" and "average" groups were sufficiently high at each age to indicate 99 or more chances in 100 that the differences found would hold in another sample of the same population The corresponding data for the boys showed no significant difference

8 Thus both methods of analyzing the data revealed relationships between IQ and height and weight for girls but not for boys during the period from three to five years It has been shown that the results cannot be accounted for on the basis of differential selection of the sex groups, greater variability of the girls, or differences in the relative maturity of the two sexes

The results obtained in this study are not in agreement with the general body of data in the field in that previous investigators do not report a sex difference in the relationship of IQ to height and weight Since, however, only meager data are available on the problem at the preschool years, it is possible that the present results may be substantiated by future investigators It must be remembered, however, that the Inquiry group is generally superior and in no sense can be considered a random sampling of the population The conclusions arrived at may therefore hold only for a comparable group On the other hand, if a sex difference in these relationships actually exists, samples of a preschool population less homogeneous than the present one may reveal differences even more pronounced than those obtained in this investigation

References

1 Bayley, Nancy Some aspects of physical growth in young children *Psychol Bull*, 1935, 32, 526-527

2 Davenport, C B Human metamorphosis *Amer J Phys Anthropol*, 1926, 9, 205-232

3. GOODENOUGH, F. L., & ANDERSON, J. E. Experimental Child Study. New York: Century, 1931. Pp. 546.
4. HONZIK, MARJORIE P., & JONES, H. E. Mental-physical relationships during the preschool period. *J. Exper. Educ.*, 1937, 6, 139-146.
5. JONES, H. E. Relationships in physical and mental development. *Rev. Educ. Res.*, 1933, 3, 150-162, 177-181.
6. ———. Relationships in physical and mental development. *Rev. Educ. Res.*, 1936, 6, 102-123.
7. PATERSON, D. Physique and Intellect. New York: Century, 1930. Pp. 304.
8. SIMMONS, K., & TODD, T. W. Growth of well children: Analysis of stature and weight 3 months to 13 years. *Growth*, 1938, 2, 93-134.

School of Medicine
Western Reserve University
2109 Adelbert Road
Cleveland, Ohio

The Journal of Genetic Psychology, 1940, 57, 83-91

STUDIES IN ANIMISM IV. AN INVESTIGATION OF CONCEPTS ALLIED TO ANIMISM*

Department of Psychology, University of Nebraska

ROGER W. RUSSELL[1]

A INTRODUCTION

Recently the author has conducted (2, 3, 4) a series of studies which are concerned with the child's verbal responses to questions concerning the animate or inanimate nature of a series of test objects He has found that the development of animism follows a course which corresponds in general to the series of successive stages discovered by Piaget (1) and has drawn conclusions concerning the various manifestations of these stages.

In the light of this work the question naturally arises as to the correlation between the manifestations of animism as exhibited by responses to inquiries concerning the "living" or "dead" nature of objects and responses to questions involving ideas which are "allied" in adult thought to those of animation, such as "knowing" and "feeling" The problem may be stated specifically as follows. Is the development of animism by characteristic stages merely a function of the child's usage of the terms "living" and "dead" or is this development actually due to more inclusive ideas concerning the nature of "life" in general

The present paper is aimed at answering this question and also at presenting detailed information as to the exact nature of the developmental course of these allied concepts

Piaget approached this very problem by studying the child's ideas of "knowing" and "feeling" In this case his technique was similar to that employed in his studies of animism (1) From the results obtained classification into stages exactly analogous to the stages of animism was possible To quote from Piaget (1, p. 173):

> For children of the first stage, everything that is in anyway active is conscious, even if it be stationary In the second stage

*Received in the Editorial Office on June 10, 1939

[1]This work was conducted under the direction of Dr. Wayne Dennis while the writer was a Du Pont Research Fellow at the University of Virginia

> consciousness is only attributed to things that can move. The sun and a bicycle are conscious, a table and a stone are not. During the third stage an essential distinction is made between movement that is due to the object itself and movement that is introduced by an outside agent. Bodies that can move of their own accord, like the sun, the wind, etc., are henceforth alone held to be conscious, while objects that receive their movement from without, like bicycles, etc., are devoid of consciousness. Finally, in the fourth stage, consciousness is restricted to the animal world.

Thus we see Piaget conceiving of a series of stages in the development of concepts allied to animism similar to the stages occurring in the development of animism itself.

This does not necessarily imply that the two series develop synchronously. For, as Piaget (1, p. 194) has pointed out, ". . . There is, indeed, nothing to show that the concepts of 'life' and of 'consciousness' are completely synonymous any more than they are to an adult." An example of the lack of correspondence between the two concepts in adults is found in connection with plants, which would be considered as having at some time or other the capacity for "life" yet would hardly be considered as ever possessing the characteristics of "consciousness."

From his investigations of the problem Piaget (1, p. 204) concluded that ". . . Although only two-fifths of the children belonged to the same stage in both series, the evolution of the two notions obeys the same laws and follows the same direction." The present experiment, which has employed more standardized and objectified methods than Piaget's studies, reveals a considerably higher correspondence between the stages of animism and the stages of allied concepts, and presents added evidence for the fundamental validity of Piaget's classification and for the sequential development of the allied concepts.

B. Method

Since Piaget discovered in his data concerning "consciousness attributed to things" four successive stages which correspond exactly with the stages of animism, the method employed in the present study followed in detail the procedure used in the examination of animism previously described by Russell and Dennis (3) with the exception of the specific questions asked. In this case the questions

were concerned with the "knowing" and "feeling" of the 20 objects[2] rather than their "living" or "dead" characteristics and thus took the following forms

1. *"Know"* In this case the questions were the same for all objects
   Does the ——— know where it is? Why?
2. *"Feel"* Here the questions were varied slightly in order to make them sound more natural to the child However, the conditions under which they were administered were the same as for the animism questions
   (1) For Objects 1 to 8 inclusive
   Does the ——— feel when I touch it? (The examiner touched each object as the question was asked) Why?
   (2) *For Object 9*
   Does the river feel when it touches something? Why?
   (3) For Object 10
   Do the clouds feel when they touch each other? Why?
   (4) For Object 11
   Does the moon feel when it touches the clouds? Why?
   (5) For Objects 12 to 20 inclusive
   Does the ——— feel when it touches something? Why?

Each subject was questioned individually, the questions involved in the present work being asked immediately following an examination for stage of animism. The questions for "knowing" and "feeling" were presented to alternate subjects so that half of the 335 subjects answered one series and half the other

Classification of the individual subjects was accomplished independently by two judges on the same basis as the classification of animistic responses (3). There was an agreement between the judgments of 99 per cent Only three cases in the 335 showed disagreement and these were classified by mutual accord after a discussion of the possible interpretations of the responses

[2]The objects referred to in this study were the same as those previously described by Russell and Dennis (3) and consisted of (1) a smooth stone, (2) a knife, (3) a mirror, (4) a broken button, (5) a comb, (6) a chair, (7) a broken dish, (8) a watch, (9) a river, (10) the clouds, (11) the moon, (12) the wind, (13) lightning, (14) a pencil, (15) a dog, (16) a bird, (17) a bug, (18) a tree, (19) a flower, and (20) the grass

C. Subjects

The 335 subjects of this study were those pupils of the "Mc" school in Albemarle County, Virginia, who were re-examined by the writer (2) in his research on the development of animism. They made up the first to seventh grades (inclusive) of this school, which serves the suburban and rural areas surrounding the city of Charlottesville. The median *IQ* for the group was 94.08, with a range of *IQ* values from 52 to 131.

D. Results

1. *Classification into Stages*

Of the 335 subjects examined only three had systematic bases for the allied concepts which were not in accord with any of the four stages. In other words, of all the subjects having some systematic basis for their responses to the examination questions, only one per cent of them are referred to as having a "Special Concept" (abbreviated *SC*), corresponding to a similar classification in the case of animistic responses.

2. *Correlation between Stages of Animism and Stages of the Allied Concepts*

It will be recalled that Piaget found only a 40 per cent correspondence between the stages of "life," or animism, and the stages of the allied concepts. That is to say, only two-fifths of his subjects were in the same stage in both concepts. The present study, on the other hand, reveals a correspondence of 63 per cent between the two series.

Table 1 presents the Coefficients of Mean Square Contingency for the correlation of stages of animism with (*a*) stages of the con-

TABLE 1

Coefficients of Mean Square Contingency for the Relationship of the Stages of Animism and the Stages of the Various Groups of Allied Concepts

Group	"C"
"Knowing"	0.76
"Feeling"	0.73
Combined	0.75

cept of "knowing," (*b*) stages of the concept of "feeling," and (*c*) stages of the combined groups[3] In all these instances the correlation involves a 6 x 6 contingency table and the coefficients cannot be expected to exceed 0 91 (5, p 66)

The fact that the correlation between animism and the allied concepts is not a perfect positive one suggests that there must have occurred many inconsistencies in the subjects' responses to the questions in the two series To the adult such inconsistencies, as, for example, Stage 2 reasons for animistic responses and Stage 3 reasons for responses to the allied concepts are immediately apparent. On the other hand, this does not seem to hold in the instances of those subjects making "inconsistent" responses In several cases where apparent inconsistency occurred the examiner took occasion to return to the animism questions after having completed inquiries on the allied concepts In all but one of these instances the subjects gave answers identical in stage with their original responses without giving any sign of recognizing inconsistencies in their explanations These facts bring up a problem of considerable interest to future research but quite outside the scope of the present paper

3 *Correlation of Stage with MA and CA*

The coefficient of Mean Square Contingency for the mental age and stage of allied concepts is 0 57, and for chronological age and stage, 0 50 In both cases the coefficients indicate a decided tendency toward a group relationship, but reveal the impossibility of individual prediction

It is interesting to note how closely the coefficients for mental age and stage correspond for animism and the allied concepts, the former being 0 59 and the latter, 0 57. On the other hand, there is some difference between the coefficients for chronological age and stage in the two concepts, that for animism being 0 62 and for the allied concepts, 0 50 It is a question, however, whether or not the latter difference is significant in the range of correlation coefficients involved.

4 *Progression of Stages*

Figures 1 and 2 and Tables 2 and 3 show the per cent of cases in each stage of the allied concepts for two-year periods ranging

[3]By "combined groups" the writer refers in this paper to the "knowing" and "feeling" groups taken as one

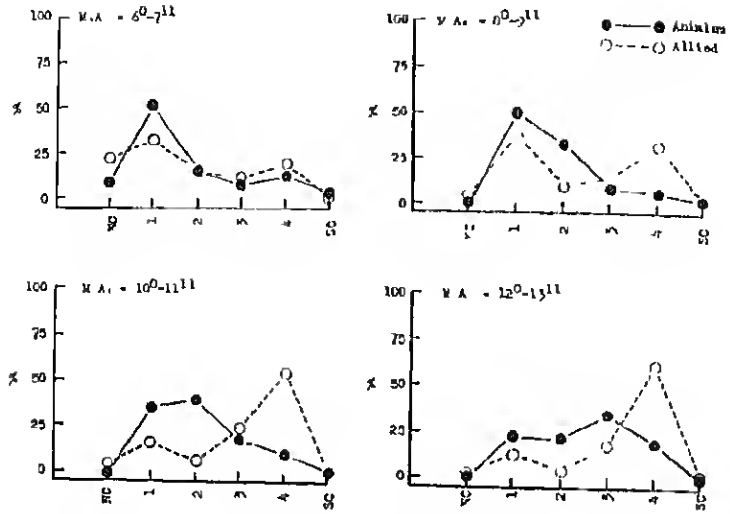

FIGURE 1
Percentage of Cases in Each Stage of Concept for Two-Year *MA* Periods (Combined Groups)

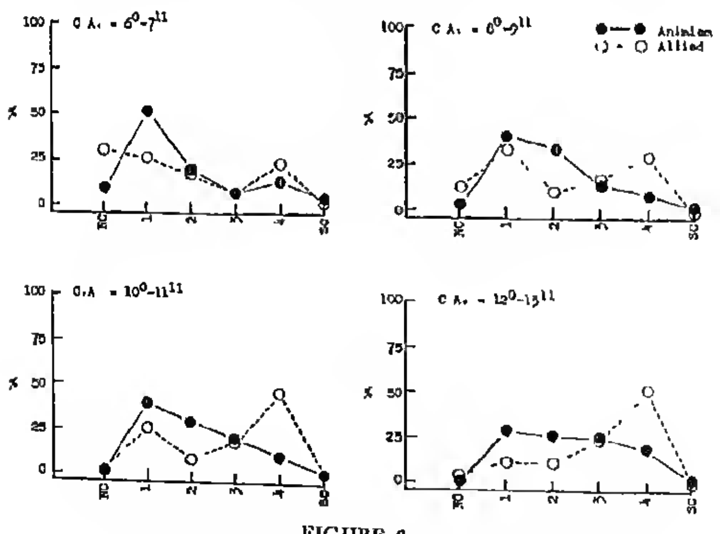

FIGURE 2
Percentage of Cases in Each Stage of Concept for Two-Year *CA* Periods (Combined Groups)

TABLE 2

PER CENT OF CASES IN EACH STAGE OF CONCEPT FOR TWO-YEAR *MA* LEVELS—COMBINED GROUPS

MA	*NC*	1	2	3	4	*SC*
6^0-7^{11}	21	32	15	11	19	2
8^0-9^{11}	4	37	10	15	32	2
10^0-11^{11}	3	15	5	23	54	0
12^0-13^{11}	2	12	4	18	62	2

TABLE 3

PER CENT OF CASES IN EACH STAGE OF CONCEPT FOR TWO-YEAR *CA* LEVELS—COMBINED GROUPS

CA	*NC*	1	2	3	4	*SC*
6^0-7^{11}	30	25	16	6	22	1
8^0-9^{11}	12	33	10	16	29	0
10^0-11^{11}	2	25	8	19	45	1
12^0-13^{11}	4	11	11	24	50	0

from mental and chronological ages of 6 years 0 months to 12 years 11 months From an examination of Figure 1 it is apparent that there is a systematic difference between the progression of the animism stages and the stages of the allied concepts which is in the direction of a greater per cent of cases of the allied concepts in Stage 4 above the 6^0-7^{11} level. The Critical Ratios for these differences at the various mental ages are as follows 6^0-7^{11}, 1 05, 8^0-9^{11}, 4 36, 10^0-11^{11}, 8 01, and 12^0-13^{11}, 5 44 These differences in favor of the allied concepts as more advanced are accompanied by corresponding differences at the lower concept stages which indicate significantly smaller percentages of allied concepts cases than of animism cases At the 6^0-7^{11} level none of the differences that occur are reliable, the nearest approach to reliability being at the Stage 2 level where the Critical Ratio was found to be 2 73

The significant differences between the stages of animism and the stages of the allied concepts are also clearly apparent throughout the chronological age range above the 6^0-7^{11} level (Figure 2) The relative percentages show again that, as a group, the subjects are more advanced in the allied concepts than in the animistic concept at all age levels above 6^0-7^{11}

These facts favoring a greater advancement in allied concept stage than in animistic stage at the higher age levels are in direct contradiction to Piaget's results Piaget says (1, p 205)

". . . children who are in the first or second stage when speaking of consciousness are generally found to be in a more advanced stage for ideas concerning life. The elder children, on the contrary, that is to say, those in the third and fourth stages, are usually in the same stage in the two parallel series."

Figures 1 and 2 and Tables 2 and 3 give ample evidence for this contradiction between Piaget's results and the present findings. However, Table 4 bears more directly on the problem. This table

TABLE 4

Per Cent of Cases in the Various Allied Concepts Stages (*a*) at the Same Stage of Animistic Concept, (*b*) at a More Advanced Stage of Animism, and (*c*) at a Less Advanced Stage

Stage	(*a*)	(*b*)	(*c*)
NC	42	58	0
1	99	1	0
2	75	3	22
3	59	1	39
4	44	0	56

shows the per cent of cases in the various allied concepts stages (*a*) at the same stage of animistic concept, (*b*) at a more advanced stage of animistic concept, and (*c*) at a less advanced stage. At the No Concept Stage level 42 per cent of the cases are at the same stage in the two series; 58 per cent are more advanced in animism than in the allied concepts, and, as would be expected since there is no lower level of classification than the No Concept Stage, zero percentage is at a less advanced level. Starting with the first stage, however, the situation is reversed. In Stage 1, 99 per cent of the cases are at the same level in both series and from then on there is a regular decrease in the per cent of cases at the same level and an increase in the per cent of cases at a less advanced level of animism with advancing concept stage, while the per cent of cases at a more advanced animistic level remains at a negligible amount. This is, then, in direct contrast to the situation as represented by Piaget.

E. Conclusions

1. *It is possible to classify individuals into the stages of concepts suggested by Piaget.* This is substantiated by the fact that 99 per cent of the subjects were classifiable, on the basis of their responses, into one or other of the designated stages.

2. *The correlation between stages of animism and stages of the allied concepts is considerably higher than the correlations of either series with mental and chronological age.* These facts answer the question raised in the introduction of this paper as to whether or not the development of animism is merely a function of the subject's usage of the terms "living" and "dead." Since in a majority of cases the subject attributes "knowing" and "feeling" to those objects that he considers as "living," and since the development of the two series of concepts follows the same progression of stages, the author feels justified in the use of the term "animism" as descriptive of the subject's ideas of "life."

3. *The correlations of the stage of concept with mental and chronological age indicate a tendency toward a group relationship, but reveal the impossibility of individual prediction.* The coefficients in this case are of the same magnitude as the corresponding coefficients for the animistic concept, being 0.57 and 0.50 respectively.

4. *Comparisons of the progression of the stages for the allied concepts and the animistic concepts reveal reliable differences which favor the allied concepts as more advanced above the 6[0]-7[11] mental and chronological age levels.* This is in direct contradiction to Piaget's results which indicate that subjects in the lower stages of the allied concepts are generally more advanced in animistic stage while older subjects are usually in the same stage in the two series. It is a question as to whether or not this is a real difference between French-speaking and English-speaking subjects, since Piaget's data are not available in statistical form.

References

1. Piaget, J. The Child's Conception of the World. New York: Harcourt, Brace, 1929. Pp. ix+397.
2. Russell, R. W. Studies in animism: II. The development of animism. *J. Genet. Psychol.*, 1940, 56, 353-366.
3. Russell, R. W., & Dennis, W. Studies in animism: I. A standardized procedure for the investigation of animism. *J. Genet. Psychol.*, 1939, 55, 389-400.
4. Russell, R. W., Dennis, W., & Ash, F. E. Studies in animism: III. Animism in feeble-minded subjects. *J. Genet. Psychol.*, 1940, 57, 57-63.
5. Yule, G. U. An Introduction to the Theory of Statistics. London: Griffin, 1919. Pp. xiii+570.

Department of Psychology
University of Nebraska
Lincoln, Nebraska

The Journal of Genetic Psychology, 1940, 57, 93-100

PERCEPTION OF SPATIAL RELATIONSHIP IN MENTALLY DEFICIENT CHILDREN*

Wayne County Training School, Northville, Michigan

HEINZ WERNER
(With the collaboration of Hilda Abrams)

One of the main problems in the development of the perception of space relations is the analysis of the perceptual cues characteristic of each genetic stage. The present study attempts to contribute to this problem as it pertains to mentally defective children

A Apparatus and Procedure

The procedure used in this experiment resembles that of the Knox Cube test[1] The child is seated before a black screen, 24"x16", containing four opaque glass windows, 1"x1" Each of these can be lighted independently by means of a bulb set at the end of a metal tube four inches long (Figure 1)

In the original Knox Cube Test the child is asked to tap four cubes in the order presented by the examiner The complete test consists of 12 combinations Our own experiments consist of two such series of 12 combinations each The first—which we shall call the "tap-series"—is very similar to the Knox Cube test The experimenter taps the (unlighted) white squares of the screen in a given order, the child is asked to repeat the tapping In a second series—the "flash-series"—the combinations are the same but are presented by successively lighted squares

The following 12 combinations were used 1-2-3-4, 1-2-3-4-3, 1-2-3-4-2, 1-3-2-4, 1-4-3-2, 1-4-2-3, 1-3-2-4-3, 1-4-3-2-4, 1-3-1-2-4, 1-4-3-1-2-4, 1-3-2-4-1-3, 1-4-2-3-4-1

B Test Results in Relation to Mental Age

It will be noted that, although the combinations of the two series are identical, the space relations are presented in the "tap-series" through concrete visual movement, and in the "flash-series"

*Received in the Editorial Office on June 17, 1939

[1]The experimental set up is based on an idea suggested by Z. Pauline Hoakley, Wayne County Training School The apparatus was built under the direction of L. C Sullivan in the shops of the Wayne County Training School

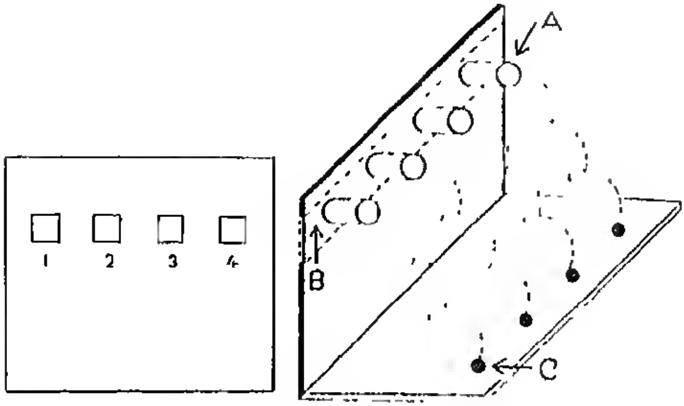

FIGURE 1

Apparatus for Administering Flash Test

Left Front View Black screen with windows 1, 2, 3, 4
Right Back View *A* Tubes containing the bulbs
B Opaque Glass
C Contacts

through relatively abstract and static optical stimuli. The main problem of this investigation is to analyze the effect of such difference in presentation at various stages of mental growth

The tests have been administered to six mental age groups of high grade moron and border line defective children of the Wayne County Training School. Each group consisted of 23-30 boys The six mental age levels are: 6 years to 6 years, 11 months, 7 years to 7 years, 11 months; 8 years to 8 years, 11 months, 9 years to 9 years, 11 months; 10 years to 10 years, 11 months, 11 years to 12 years

In order to avoid any practice effect the two series were administered approximately three weeks apart. The results are shown in the graph, Figure 2 The curves present the average error scores of the successive mental age groups, for each series Since the standard scoring in the Knox Cube test did not seem to us to give a fair picture of individual achievements, the following more differential method of scoring was used:

The interchange of two neighboring numbers (e.g., 1-3-4-2, instead of 1-4-3-2) was counted one error point The omission of one number (e.g., 1-4-3, instead of 1-4-3-2) was also scored as one error point The addition of one number (e.g., 1-3-4-3-2 instead of

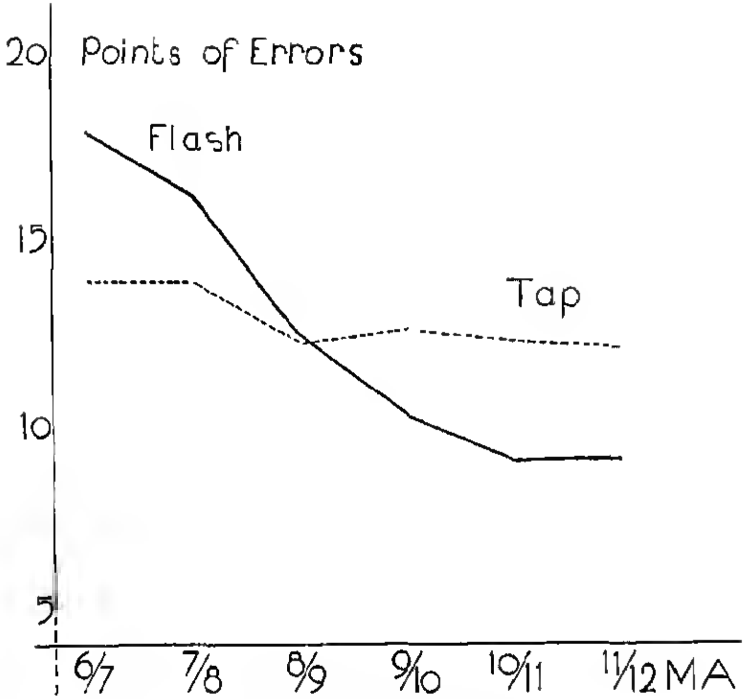

FIGURE 2
Error Curves for Flash and
Tap Test in Relation to Mental Age

1-4-3-2) was scored as two error points Any other error received the maximum score of three error points

The graph shows a decrease of errors with increasing mental age The decrease is present in the tap- as well as the flash-series, although it is more consistent and much more marked in the latter On the tap test the total decrease is only 13 per cent, in contrast to almost 50 per cent on the flash test This fact implies a striking genetic change in the relation between the achievement on the two tests. Whereas the lower mental age groups are superior on the tap test, the highest mental age groups score higher on the flash test The reversal in the relation of the two performances seems to occur between the mental ages of eight and nine years This con-

clusion is supported by the fact that at the 8/9 year level the critical ratio of the difference of the means is low (0.6), while at all other levels it approaches or reaches statistical significance (between 2.8 and 3.5, as indicated in Table 1).

TABLE 1
RELATION OF TEST SCORE TO MENTAL AGE

MA	Flash Error points	Flash *SD*	Tap Error points	Tap *SD*	*Diff* (Flash-Tap)	$\frac{Diff}{\sigma Diff}$
6/7	17.8	5.2	13.8	3.5	+4.0	3.3
7/8	16.0	3.0	13.8	3.0	+2.2	2.8
8/9	12.4	3.5	12.2	4.1	+0.2	0.6
9/10	10.2	2.9	12.5	3.1	—2.3	2.8
10/11	9.0	3.0	12.2	4.0	—3.2	3.2
11/12	9.0	3.1	12.0	3.0	—3.0	3.5

Another question arises at this point. Is this reversal in the test results actually due to the increasing growth in mental age? Or is it the result of the difference in *IQ*, that is, in relative intellectual capacity, between the lower and the higher age groups? Naturally, in a random selection of mental age groups from a population of mentally deficient children, higher mental age groups tend to have higher *IQ*'s. This occurred in our selection. The *IQ*'s of the six age levels are as follows: six year level: 60, seven year: 62; eight year: 67, nine year: 70, ten year: 73, eleven year: 77.

In order to answer the above question we selected from the total group the children whose *IQ* lay between 63 and 73. The average test results of these children, grouped according to the six mental age levels, have been computed in Table 2. Although the groups vary considerably in size we may conclude from the table that the general trend found in the larger groups is characteristic also of the groups within a restricted *IQ* range. The findings suggest, therefore, that the growing superiority in achievement on the flash test over the tap test is actually a function of mental growth.

C. Discussion

The present study supports the view of those psychologists who believe that development in the perception of space relations is not a purely quantitative affair. Genesis of space relationship not only involves greater and greater complexity but has a definite qualitative

TABLE 2
RELATION OF TEST SCORE TO MENTAL AGE WITHIN A RESTRICTED *IQ* RANGE (63-73)

	Flash		Tap	
MA	Error points	*SD*	Error points	*SD*
6/7	15	4.2	12.5	3.2
7/8	14.5	3.5	12.5	3.1
8/9	13.5	3.8	11.7	2.9
9/10	10.8	2.5	12.0	3.0
10/11	8.8	3.2	11.5	3.2
11/12	8.5	2.5	11.5	2.8

aspect, since, during growth, a change of dominance occurs among certain factors underlying the perception of space. This shift of dominance is indicated, in our experiments, by the growing superiority of the performance on the flash series over the achievement on the tap test.

The psychological differences of the perception of space relationship in the two series seem to be the following: (*a*) Spatial relationship as perceived in the tapping is rather concrete and personal; spatial relationship perceived in the flash technique is more abstract and impersonal. (*b*) In the tap series, visual movements are outstanding characteristics of the stimuli, whereas flashes are more static in nature.

According to normal genetic psychology space perception develops in terms of increasing objectivation and abstraction on the one side, and of an increase of static in relation to dynamic characteristics (4) on the other. These ontogenetic principles have been based primarily on careful observations (3), and to a lesser extent upon experiments. Two experiments may be cited as evidence of this development.

The experiments of *Meili* and *Tobler* indicate that the visual spatial field of the younger child is dynamic to a greater degree than that of the older child and of the adult. The authors compared the threshold of apparent movement in 38 children of 5 to 12 years of age with that of 22 adults (1). The younger children could discern movement in the stroboscopically projected (cinematic) images at a lower speed than could the older children and the adults. In other words younger children are more sensitive to apparent movement than older children.

A second experiment concerns the development from concrete, personal to abstract, objective space relation. Piaget devised a group

of tests which are very useful in the analysis of this development According to Piaget's findings, a child is able relatively early to distinguish left and right on other persons, but learns to discriminate left and right on inanimate objects at a much later time For instance, the spatial relations of three different objects placed side by side cannot be conceived correctly until the age of 10 to 11 years (2).

Our own experiments agree with these findings rather well They indicate that the structure of visual space undergoes a transformation with increasing mental age Objective, quasi-geometric properties of space and space relations develop rather rapidly at the age levels concerned, becoming more important than the more primitive, concrete, and dynamic properties of spatial relationship

D THE SPACE RELATIONS TEST AND SPECIAL ARITHMETIC DISABILITIES

In a previous paper a definite relationship has been demonstrated between special arithmetic disability and low scores in certain visual tests (5) In these tests a group of 14 boys high in arithmetic achievement proved superior to a similar group low in arithmetic achievement, in their ability to reproduce various patterns composed of dots. The average chronological age, mental age, arithmetic and reading age of the two groups are indicated in Table 3

TABLE 3

	CA	*MA*	Read *A*	Arith *A*	
Poor arithmetic group	150	102	105	95	Months
Good arithmetic group	160	105	98	112	Months

The present test has been used for further analysis of the visual factors related to arithmetic disability. The test was administered to the two groups. Their scores are indicated in Table 4. The results suggest that the poor arithmetic group is definitely inferior

TABLE 4

	Points of error Flash	Tap
Poor arithmetic group	14 5	11 3
Good arithmetic group	9 5	11.7

on the flash test with respect to the tap test. The performance of the good group on the flash test is definitely superior to their performance on the tap test The conclusion that the score on the flash test is related to arithmetic achievement has been strengthened by calculating the deviation of each individual's score from the average score for his mental age Fifty-eight per cent of the poor group have a lower performance on the flash test than the average for their mental age, for 22 per cent the score is equal, 22 per cent exceed the average Fourteen per cent of the boys of the good group score lower, and 58 per cent higher than the average for their mental ages, 28 per cent score at the average [2]

We may therefore conclude that the ability to grasp quasi-abstract visual space relations seems to have a definite bearing on achievement in arithmetic [3]

E Summary

1 Approximately 180 moron and borderline defective boys 6 to 12 years of mental age were given a test which resembles the Knox Cube test The present test consists of two experimental series In Series I the experimenter taps four squares in a given order, and in Series II, the same combinations are presented by lighting the squares In both series, the child is requested to tap the squares in the prescribed order.

2. In both series, the errors decrease with increasing mental age

3. The performance on Series II is inferior to that of Series I at the lower mental age levels, superior at the higher mental age levels

[2]The difference in the two scores may be extreme in individual cases Such a case is J B, a boy with a severe arithmetic disability. The case has been described in previous publication by A. Strauss and myself (Finger agnosia in children *Amer J Psychiat*, 95, 6, 1939.) The boy, with a mental age of 9 years, 10 months is retarded 27 months in arithmetic as compared with reading His error score on the tap series is only 7, but he scores 20 error points on the flash test

[3]Further analysis of the data has revealed differences in the performance of children belonging to two clinical types of mental deficiency, viz, the endogenous (heredofamilial) and the exogenous type (characterized by brain lesion) It has been shown previously that these clinical types perform visual-motor tasks in characteristically different ways (Cf, H Werner and A A Strauss Types of visuo-motor activity in their relation to low and high performance ages *Proc Amer Ass'n Men Defic*, 1939) Similar qualitative differences in the performance on the present test have been found, and will be described in a later publication.

4. In mentally deficient children, scores on the flash series seem to vary much more widely with mental age than on the tap series. It may therefore be inferred that the flash series is more discriminative than either Series I or the original Knox Cube test.

5. There seems to be a definite relation between special arithmetic disability and poor performance on the flash test (not on the tap test).

References

1. Meili, R., & Tobler, E. Les Mouvements Stroboscopiques chez Les Enfants. *Arch. de Psychol.*, 1931, **23**, 131-156.
2. Piaget, J. Judgment and Reasoning in the Child. New York: Harcourt, Brace, 1928.
3. Stern, W. Psychology of Early Childhood. New York: Holt, 1930.
4. Werner, H. Comparative Psychology of Mental Development. New York: Harper, 1940.
5. Werner, H., & Strauss, A. Problems and methods of functional analysis in mentally deficient children. *J. Abn. & Soc. Psychol.*, 1939, **34**, 37-62.

Wayne County Training School
Northville, Michigan

The Journal of Genetic Psychology, 1940, 57, 101-118

CONSISTENCY OF FOUR METHODS OF MEASURING ONE TYPE OF SPORADIC EMOTIONAL BEHAVIOR (CRYING) IN NURSERY SCHOOL CHILDREN*

Institute of Child Welfare, University of California

CATHERINE LANDRETH[1]

A. METHOD OF INVESTIGATION

The investigator of emotional behavior in preschool children has several methods open to him. Briefly summarized these are the experimental method, its development in terms of a more precise instrumental measurement, illustrated by the galvanometric technique, controlled observation of repeated short samples of behavior, diary records taken in systematic form, ratings of behavior traits, interviews, and projective methods. With this variety, and a literature resulting from the employment of different techniques, it becomes essential to determine to what extent choice of method used in investigating a particular type of behavior may affect the results obtained.

That method may affect results is indicated by M. C. Jones and B. S. Burks (3) review of investigations on behavior traits. A tabulated summary of 10 investigations in which correlations between ratings and objective measures of children's traits were listed reveals that 15 of the 34 correlations were below .50. Robinson (4), who used four measures for investigating emotional expressiveness, reports correlations of approximately zero. How may this lack of agreement be explained? Is it a matter of crudity of the recording device and unreliability on the part of the recorder, does it arise through lack of analysis and hence lack of standardization of the situations in which the behavior was observed, or, would a record of variations in individual endogenous factors reveal relationships which would affect agreement, even when the exogenous factors remained constant from one situation to another?

*Accepted for publication by Harold E. Jones of the Editorial Board, and received in the Editorial Office on June 18, 1939.

[1]The writer is indebted to Drs. H. E. Jones, H. S. Conrad, and Mary Shirley for reading and criticism of the manuscript. Clerical and statistical assistance in the preparation of the material was furnished by the personnel of Works Progress Administration Official Project No. 65-3-5406.

The present investigation was undertaken with a view to determining the extent to which these factors might be operative in investigations of emotional behavior The behavior investigated was crying The choice arose from the need of observing this behavior in all the situations arising during the course of a day This necessitated cooperation on the part of the children's parents It was assumed that parents would be more willing to investigate and less apt to overlook a form of behavior that constituted a nuisance to them. This assumption was fully justified in the zeal with which the parents undertook cooperation.

The evidence furnished from this study included.

1 A comparison of results from four methods of investigation carried on simultaneously[2] in one situation—the free play period in a nursery school

2 A comparison of results from one method of investigation employed in two different situations—the free play period in a nursery school,—and the home of each child

3 A record of variations in exogenous factors—(children and adults present, activity engaged in)—resulting in classification of the situations in which crying occurred, and data on the incidence of crying in different situations.

4. A record of variations in endogenous factors—colds, minor illnesses, constipation, etc, and the relationship of such factors to incidence of crying behavior.

5 A record of the responses obtained from different individuals in the environment to this behavior, and the relationship between type of response and incidence of behavior

6 Case analysis revealing dynamics of interaction between these various factors

The present report is limited to a comparison of methods employed in terms of their reliability and validity A forthcoming study will present statistical findings on factors associated with crying and an analysis of the situations provoking this behavior, in terms of their significance for particular individuals, and for age and sex grouping.

Thirty-two children, 14 girls and 18 boys, with age range from two years eight months to five years two months at the midpoint of the investigation, made up the sample The children had been

[2]Ratings on the children were made two months prior to the period in which objective measures of the behavior investigated were obtained

in attendance at the Institute of Child Welfare nursery school for periods varying from 36 to 393 days Mean and median *IQ*'s were 122 and 124 respectively The group represented a selected socio-economic sample in that they were drawn from middle class homes of professional and business people The selection was necessitated by the need for obtaining intelligent cooperation from the parents in the matter of record keeping

The four methods of investigation applied in the nursery school were incident sampling over a period of eight weeks, time sampling over a period of seven weeks, the head teacher's record for the eight weeks period and the same teacher's rating made shortly before the investigation was undertaken.

The infrequent character of crying behavior, commented on by other investigators, suggested the possibility of obtaining a reliable record through observations of the total group Such a procedure was facilitated by the following conditions (*a*) the children were out doors during the entire period of observation, (*b*) practically the entire outdoor space could be surveyed from one position, (*c*) the morning program was devoted entirely to free play save for a short period of story telling and a brief interval for fruit juice

The term "incident sampling" was applied to this method of recording because it was impossible to record absolutely every incident of crying behavior However, over a period of 40 days it was undoubtedly possible to obtain a large and representative sampling of each child's behavior for the period of observation As indicated in the record form (Table 1) use of numbers in place of names for the children, code letters for equipment in the playground, and check marks for the majority of items made rapid recording possible The following criteria were furnished in regard to the various types of crying behavior

> *Screaming* varies in intensity, but is characterized by a shrill aggressive quality A prolonged outcry without tears belongs in the screaming category
>
> *Squealing* includes the short, often staccato squawks of protest It is not prolonged like screaming, nor quite as aggressive
>
> *Whining* is characterized chiefly by the nasal complaining tone Though often associated with words, it also accompanies non-verbal vocalization.
>
> *Sobbing* usually occurs after crying, it is characterized by convulsive breathing There may still be some tears, but it is distinguished from crying in not including a vocal outcry.

TABLE 1
OBSERVATIONAL STUDY OF CRYING

Date	Day						Recorder						
Children present		1	2	3	4	5	6	7	8	9	10	11	12
		13	14	15	16	17	18	19	20	21	22	23	24
		25	26	27	28	29	30	31	32	33			
Child crying TIME	(1) Time of occurrence Duration of outburst												
SITUATION	(2) What "set off" crying? (3) *Who were present?* What was child doing at the time?												
CRYING BEHAVIOR	(4) Crying with tears Screaming Sobbing Whimpering Whining Squealing												
ACCOMPANYING BEHAVIOR	(5) Words spoken Attacking-hitting, kicking, etc Running for help—Asking help Indirect, limpness, rigidity, throws self on floor—holds breath												
REACTION OF THOSE PRESENT	(6) Teacher Children Type of response												
EFFECT ON CHILD'S BEHAVIOR	(7) Stops behavior Continues Increases												
AFTER EFFECTS	(8) Cheerful Fretful and unhappy Other												

Whimpering differs from sobbing in having a vocal element—a broken cry It is easily distinguished from crying by this broken cry However it also frequently develops *into crying*, if you are in doubt, and there are tears, check the outburst as crying with tears

A preliminary practice period furnished evidence as to the order of checking which gave greatest reliability This order, indicated by numbers on the record form, was adhered to throughout the period of observation Information furnished in the sheet of instructions included the following

The Situation

What set off crying? As much of the crying seems to come from attacks or encounters with other children, check specifically these three items—the number of the "attacker," the nature of the "attack," at what it was directed—the child's person or property What was child doing at time? Use code letters for play equipment Do not attempt to get details of child's activity, simply his location with reference to the particular group of play equipment

Type of response

Check first *by whom* Watch for response from child causing crying He may desist in his attack, redouble his efforts, or simply continue Also note any response from other children as well as adults

After Effects

Observe child for one minute after crying has ceased unless another child cries Check after effects Check as cheerful unless there is a definite negative reaction Cheerful simply means a return to the child's usual state If you are observing one crying incident and you hear another do not try to get it, put a cross in the upper right hand corner of the sheet you are recording on to show you have missed an incident Later, you may be able to check the child's letter beside the cross Do not rush through one incident to get another

At the end of the practice period, recording for the eight weeks duration of the investigation was undertaken by a graduate student in psychology, save for a two-weeks period during which the writer recorded simultaneously to obtain a measure of the reliability of the observations A lap board with a stop watch and watch securely attached, a sheaf of records, and a pencil were all the equipment required for record taking

As the record used in the time sampling or time frequency observation indicates (Table 2) the time unit was 15 seconds The length of observation was one minute repeated four times during

TABLE 2

THE EXPRESSION OF NEGATIVE EMOTIONAL REACTIONS

Child . Observer Date

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

21 22 23 24 25 26 27 28 29 30 31 32 33

	Time			
Frowning				
Scowling				
Pouting				
Crying				
Crying (Tears)				
Screaming				
Sobbing				
Whimpering				
Whining				
Squealing				
Talking				
Kicking				
Striking				
Pinching				
Scratching				
Pulling hair				
Biting				
Running at				
Pulling				
Pushing				
Shrinking				
Struggling				
Running for help				
Running away				
Jumping up and down				
Throwing self on F				
Stiffening				
Limpness				
Refusing to move				
Throwing things				
Nervous habits				

Each short horizontal line represents a fifteen-second interval

each morning, the children being observed in serial rotation One hundred or more one-minute observations were obtained on 23 of the 30 children observed. Absences made it impossible to secure this number on the remaining seven Absences also made it necessary to take more than four one-minute records a morning on seven children During the last three days of observation eight one-minute records were obtained on these children As it was desired to compare the results obtained from these records with those of other investigators who had observed only during free play situations, no records were taken during the story period or while a child was getting his orange juice or going indoors to the bathroom

During the entire period of the investigation the head teacher recorded at the end of the morning the names of all the children whom she recalled as having cried The reason for obtaining this type of record which was necessarily a much cruder measure of the child's behavior than the two measures previously described was primarily for purposes of comparison with the head teacher's rating At the beginning of the study this teacher had rated 25 of the children on 61 items from the California Behavior Inventory for Nursery School Children (2) As one of the items in the scale was the inhibition of crying, it was believed that the same teacher's record and rating on this behavior trait might furnish some evidence as to the extent to which teachers' ratings are based on a mental photographic record of the child's actual behavior While it is true that the teacher's rating applied to a period before the study was made, the length of time elapsing between the actual rating and the conclusion of the investigation was only two months

The type of record kept of the child's crying behavior at home for a period of five weeks was almost identical with the incident sampling form used in the nursery school The only modifications were that numbers for children present were omitted, the section *Time* included an item "reported by", and the section *Reaction of those present* had the item "by whom" in place of "teachers, children" In addition to this record form, parents furnished information on a systematized daily record sheet as to the child's physical condition and health routine his sleeping and nap hours, elimination, meal hours, and toilet accidents They also recorded events varying the day's routine such as adult or child visitors in the home, excursions or visits made by the child outside the home The sheet of instructions furnished the parents is printed below

1 Records are to be left in the box on the nursery school mantel-piece when the children are brought to school in the morning (On Monday mornings there will be three records one for Friday, Saturday, and Sunday)

2 In the space marked "*Hours not included,*" at the foot *of the first sheet, record is to be made of any periods during* which the child was not under observation (e.g, "mother away shopping") Record the beginning and the end of the period (e g, from 2 30 to 4 40 p m) Place a check mark in this space if child was under observation during the entire day

3 Try to record every incident of crying with tears, or screaming, or sobbing (convulsive catching of the breath), or whining, or whimpering (broken crying), or squealing

After using the blank for a few days the time required for recording will be greatly shortened For the most part the *record can be made by placing a check mark opposite the* appropriate entry If one sheet is insufficient for recording the number of incidents occurring during the day, the second sheet of other blanks may be used and pinned or clipped to the record

4 In recording duration of outburst use

	if outburst is more or less momentary			
1	if outburst occupied less than 2 minutes			
5	"	"	"	from 2 to 5 minutes
10	"	"	"	from 5 to 10 minutes

If longer than 10 minutes state time as nearly as possible

5 It is desirable to have all records made by the same person as far as this is possible, but in cases where the child is cared for by several persons at different times, it may be necessary for a number of people to participate in the study in order to insure a complete report Opposite the space marked "reported by" write the name or initials of the person reporting the particular incident or crying or related behavior

6. (*A*) "*What was child doing at the time*" and

(*B*) "*What set off crying*" should be recorded as briefly and concisely as possible and yet give an adequate picture of the situation (e g, (*A*) Playing with her doll; (*B*) Brother pulled it away from her)

7 Opposite "*Type of response of those present*" report exactly what was said or done, and in the space "*By whom*" record the name or initials of person making the response

8 In the section "*Effect on Child's Behavior of this Response*" put two checks (√√) opposite "*stops behavior*" if child stopped crying or screaming, etc, immediately, one check (√) if crying or screaming tapers off slowly"

A question naturally arises as to the accuracy and significance of results obtained in this way from untrained observers. As 60 per cent of the mothers were college graduates and the majority of the remaining 40 per cent had one or more years of college education,

TABLE 3

SUMMARY OF MAIN DATA ON INDIVIDUAL CASES

					Incident sampling					
					Home			School		
Case No	Sex	Age Yrs	Age Mos	IQ	Waking hrs recorded	Crying incidents per day	Time spent crying	Hours recorded	Crying incidents per day	Time spent crying
					no	no	%	no	no	%
1	F	1	0	115				99	2.00	0.40
2	M	4	8	132				96	1.03	0.34
4	M	4	0	137	324	2.66	1.09	114	3.39	1.09
5	F	4	0	116	306	0.71	0.03	117	1.44	0.33
6	M	2	9	124	347	0.91	0.17	87	0.03	0.00
7	M	3	7	127	299	1.06	0.94	51	0.65	0.26
8	M	2	8	116	310	1.06	0.15	66	0.41	0.13
9	M	5	0	113	253	0.97	0.07	90	0.57	0.16
10	M	5	2	117	318	0.31	0.20	117	0.03	0.00
11	F	3	2	125				81	1.04	0.25
12	F	3	2	121				81	0.67	0.15
13	M	3	2	135	311	2.37	1.11	87	1.14	0.42
14	M	3	5	141	282	2.14	0.14	78	0.12	0.02
15	F	3	10	124	271	0.77	0.17	99	0.18	0.05
16	M	3	4	138	294	1.06	0.31	60	1.45	0.20
17	M	2	10	95	229	1.04	0.39	108	0.33	0.17
18	F	4	10	128	319	0.17	0.08	99	0.58	0.19
19	M	4	6	102	274	2.07	0.79	111	1.38	0.45
20	F	2	10	120				111	2.27	0.46
21	M	3	9	101	261	1.17	1.18	99	5.69	1.02
22	F	3	6	122	341	0.40	0.60	105	0.77	0.19
23	M	3	9	104	314	1.62	0.21	105	0.34	0.09
24	F	1	10	132	265	0.31	0.04	66	1.50	0.25
25	F	4	4	136	301	3.06	0.73	120	0.98	0.22
26	F	3	9	129	354	0.91	0.20	72	1.08	0.19
27	F	3	4					75	0.92	0.16
28	M	4	3	123	311	1.31	0.30	87	0.72	0.14
29	M	3	3	86	264	0.97	2.02	102	2.97	0.98
30	F	3	5		302	0.54	0.28	105	0.86	0.38
31	M	3	6	116	295	1.43	0.34	96	1.44	0.36
32	F	2	10					30	4.20	0.63
33	M	4	7	130	329	0.97	0.10	93	1.58	0.35
					Mean	1.19	.48		1.30	.31
					Median	1.01	.30		1.04	.21

TABLE 3 (*continued*)

					Time sampling School		Teachers' record School		Teachers' rating
Case No.	Sex	Age Yrs.	Age Mos	*IQ*	Minutes observed	Time spent crying	Days recorded	Days crying occurred	7-Pt Scale
					no	%	no	%	
1	F	3	0	115	108	1.16	33	85	5
2	M	4	8	132	108	1.85	32	34	5
4	M	4	0	137	120	2.71	38	82	
5	F	4	0	116	125	0.80	39	49	4
6	M	2	9	124	108	0.00	29	00	3
7	M	3	7	127	71	5.63	17	35	
8	M	2	8	116	61	0.00	22	36	3
9	M	5	0	143	100	0.50	30	10	3
10	M	5	2	117	125	0.20	39	5	2
11	F	3	2	125	100	0.75	27	48	4
12	F	3	2	121	100	1.50	27	18	4
13	M	3	2	135	103	3.16	29	52	4
14	M	3	5	141	79	0.00	26	12	
15	F	3	10	124	103	0.49	33	10	3
16	M	3	4	138	54	0.00	20	45	4
17	M	2	10	95	113	1.77	36	41	5
18	F	4	10	128	105	1.19	33	33	4
19	M	4	6	102	115	3.26	37	76	7
20	F	2	10	120	119	0.21	37	76	5
21	M	3	9	101	105	5.00	33	97	6
22	F	3	6	122	110	0.23	35	31	2
23	M	3	9	104	105	0.48	35	29	4
24	F	4	10	132	68	1.10	22	45	4
25	F	4	4	136	128	1.56	40	50	4
26	F	3	9	129	92	1.90	24	46	4
27	F	3	4		.				
28	M	4	3	123	93	2.15	29	45	3
29	M	3	3	86	107	2.57	34	85	5
30	F	3	5	.	105	0.71	35	43	
31	M	3	6	116	114	2.41	32	59	5
32	F	2	10	.					.
33	M	4	7	130	101	2.23	31	73	
					Mean	1.5		42	
					Median	1.34		45	

it is reasonable to assume that they were equal in education and ability to the average of the technical assistants employed in most research laboratories. Further, the records were phrased in such a way as to call for a minimum of interpretation and descriptive recording.

Evidence as to the possibility of the mothers' modifying the records

in order to place their own or their child's behavior in a favorable light would seem to be lacking in the findings which will be discussed later While it is possible that some mothers did not record every incident of crying that occurred during the 35 days of record taking, it is safe to assume that each mother obtained a large and representative sampling of her child's crying behavior during that period.

Not all of the records were made on 35 consecutive days In some instances the mothers missed one day or more when the child was at home sick. This was due to two facts. one, that the child cried much more when sick, the other, that the additional time called for in giving him adequate physical care made record keeping an impossibility for the time being

Table 3 summarizes the data obtained from methods employed

B Results

1 *Reliability of Observers*

a Percentage agreement between observers using time sampling method The number of one-minute observations on which agreement between two observers was based was 691. As the time unit used was 15 seconds, this made 2764 periods in which agreement was possible During 2004 of the 2764 periods both observers agreed in recording an absence of any of the behavior listed in the record forms. Using Arrington's method (1) *the percentage agreement between observers for periods when behavior was recorded was 97.5 per cent.*

b. Percentage agreement between observers using incident sampling method. Two observers recording simultaneously and independently over a period of 10 days obtained a record of 285 incidents An analysis of the percentage agreement calculated as

$$\frac{\textit{Total No. incidents—Total instances of disagreement}}{\textit{Total No incidents}}$$

gave the following figures

1	Time of occurrence (within one minute)	98 0
2	Duration of the crying incident (within 15 seconds)	97 6
3	Name of child or adult associated with crying behavior	98 8
4	The cause of crying (seven categories)	99 6
5	Play equipment which incident centered around	100 0

6.	Type of crying behavior (seven types)	98.4
7.	Type of accompanying behavior (four types)	96.8
8.	Type of response (eleven categories)	87.0
9.	Person responding (name of child or adult)	89.4
10.	After-effects (cheerful or not cheerful)	90.0
	Mean percentage agreement	95.5

As it was possible to have more than one instance of disagreement on Items 4, 6, 7, 8, and 9 for each incident, this method weights lack of agreement and gives too low a figure for these particular items. The seven types of crying and four of accompanying behavior are those listed in the record form. The lowest agreement was obtained on Items 8 and 9, type of response and person making the response. Lack of agreement on Item 8 in all save two instances was due to an omission by one of the observers. In Item 9 every instance of disagreement represented an omission. As it was possible for the child crying to obtain various types of response from children as well as adults, the accurate observation of all responses called for a high degree of visual alertness as well as swift recording. For this reason a certain amount of disagreement was practically unavoidable.

Despite these difficulties both mean agreement and agreement on separate items indicate that the incident sampling method furnished a reliable means of obtaining the type of data sought in this particular investigation.

2. *Consistency of Crying Behavior*

The consistency of the data on time sampling observations was tested by comparing the frequency of occurrence of crying in the following two ways:

1. A comparison of the first half of the total record with the second half.

2. A comparison of the first two minutes with the last two minutes of the record of each child for each of the mornings of the observation.

By the first method a correlation of .01 ± .12 was obtained; by the second, a correlation of .10 ± .02. The consistency of the data on incident sampling was tested by comparing the percentage of time spent crying by each child on odd and even days. A correlation of .24 ± .02 was obtained. This lack of consistency in crying behavior is a feature of its sporadic nature. As Table 3 indicates, the average amount of time spent crying per child was .31 per cent of the time

he spent in the nursery school, or about half a minute per day. As the time-sampling records were based on a sample of four out of the 180 minutes a child was present each morning, the difficulty of getting a representative sample of his behavior by this means would seem obvious.

The low degree of consistency found for the incident-sampling method finds explanation in the evidence furnished by the present study that crying occurs as the result of fairly specific situations, and that it is influenced by variations in the endogenous as well as the exogenous factors in the environment.

3. *Correlations Between Different Measures of Crying and Related Behavior*

The degree of relationship between incidence scores for crying obtained by different methods is summarized in Table 4. Fourteen of the 21 correlations are over .50. Of the seven below .50, six represent relationships involving a comparison of the number of incidents of crying behavior in the home with scores obtained from each of the other six methods. Apparently the number of crying incidents in the home provides the least representative record of the child's crying behavior.

The average of all the correlations between different measures is .55. This figure is indicative of the extent to which the tendency to cry is a general one. This would seem to suggest that if for any reason a child cries in one situation, he will be apt to cry in another,—as for example, at home and at school. However, these r's are not so high as to warrant the assumption that if a child cries in one situation he will cry in every other one. The low degree of consistency of crying behavior by the split-half method, suggests rather that it is not a stable personality characteristic. As will be shown later it is affected both by strain and stress in the environment and by variations in the child's physical condition. Hence the correlation between different measures indicate that for cross sections of the child's behavior, regardless of what method of recording is used, somewhat the same picture is obtained. On the other hand, in a longitudinal survey of the child's behavior, comparison between incidence of crying at different periods shows no such degree of relationship because different factors modifying his behavior are in operation. In other words the lack of consistency in the behavior is due to the lack of consistency in the factors that affect it.

TABLE 4

CORRELATIONS BETWEEN INCIDENCE SCORES OBTAINED BY DIFFERENT METHODS

	Incident-sampling N S % time	Incident-sampling N S No incidents	Teachers' record % days	Teachers' rating 7-point Scale	Incident-sampling H % time	Incident-sampling H No incidents
	r	r	r	r	r	r
Time sampling % time	.58 ± 08	55 ± 08	51 ± 09	66 ± 07	.64 ± 08	.28 ± .12
Incident-sampling N S % time		.90 ± 02	.81 ± 04	67 ± 07	78 ± 05	.34 ± 11
Incident sampling N S. No incidents			80 ± 04	58 ± 09	60 ± 08	10 ± 13
Teachers' record % days				73 ± 07	67 ± 07	30 ± 12
Teachers' rating 7-point Scale					41 ± 12	43 ± 12
Incident-sampling H % time						.28 ± 12

Considering now the variations in degree of interrelationship of each measure with every other measure, the correlation averages are as follows:

Incident sampling in the nursery school, Percentage of time spent crying	68
Teacher's record, Percentage of days crying recorded	64
Incident sampling in the nursery school, No. of incidents	59
Teacher's ratings, Seven-point scale	.58
Incident sampling in the home, *Percentage of time spent crying*	*56*
Time sampling in the nursery school, Percentage of time spent crying	54
Incident sampling in the home, No of incidents	29

Of considerable interest is the fact that *r* between number of incidents and percentage of time spent crying in school is 90± 02, while the *r* between number of incidents and percentage of time spent crying at *home* is only *28± 12* This would seem to suggest, even when allowance has been made for less accurate time keeping, that the length of time children cry at home is influenced differentially by the response to their crying In the nursery school, where all *children meet a more or less uniform response, the number of* incidents of crying is roughly indicative of the time spent crying. No such prediction can be made for crying behavior in the home Reference to Table 3 shows that child No 29, who has an average of 97 *incidents of crying a day at* home, *spent* 2 02 *per cent of his time* crying, while child No 25, with an average of 3 06 incidents, spent only 73 per cent of the time crying

The lack of relationship (10± 13) between the number of incidents recorded in the *nursery* school and the home is testimony to the extent to which specific situations are responsible for crying episodes in children. Referring to Table 3 it will be seen that child No 21, who ranked first in number of incidents of crying in the *nursery* school, ranked seventeenth in number of crying incidents at home Child No 25, who ranked first in number of incidents at home, ranked sixteenth in the number of crying outbursts in the nursery school On the other hand, when the percentage of time

child No 21 spends crying at home and at school is compared, it will be seen that his position in regard to the group is the same for each situation. He ranks second in both groups The correlation of $.78 \pm .05$ for percentage of time spent crying at home and at school would seem to indicate that the percentage of time a child cries furnishes a more significant index to his crying behavior than the number of times he cries

C Summary

This investigation had a two-fold purpose On the one hand it sought to discover the relation of various factors to one type of overt emotional reaction—crying—in preschool children, on the other, it was equally concerned with a determination of the extent to which different observational methods may influence the type of result obtained The study was based on daily observations made over a period of two months on a group of 32 children attending the Institute of Child Welfare nursery school Fourteen girls and 18 boys, of ages ranging from two years eight months to five years two months at the midpoint of the investigation, made up the group The children represented a selected sample from the socio-economic standpoint, as they were drawn from middle class homes of professional and business people The selection was necessitated by the need for obtaining intelligent cooperation from the parents in the matter of record keeping.

The methods of observation employed were (*a*) Incident sampling in the home over a period of five weeks, the records being kept by the children's mothers, (*b*) incident sampling in the nursery school over a period of eight weeks; (*c*) time sampling in the nursery school over a period of seven weeks, (*d*) head teacher's records in the nursery school over a period of eight weeks, (*e*) one teacher's ratings based on her acquaintance with the children during the period they were in attendance at the nursery school. Fairly detailed home records reporting the daily health routine of the children were also obtained for a period of five weeks The purpose of these records was to determine the relation between health routine and emotional upsets

Behavior was recorded with a high degree of reliability represented by 97.5 per cent agreement between two observers for time sampling and an average of 95 5 per cent (range 87 to 100) for 10 items recorded by incident sampling

Incident sampling furnished the most valid measure for sporadic behavior of the type investigated As will be discussed in a later article it also furnished reliable evidence as to the situations which called forth this behavior in different children

Time sampling furnished a much less satisfactory and less valid record As a method of investigation it is obviously unsuitable for sporadic types of behavior, because the length of observation—four minutes—is too short to obtain a representative record of behavior which occurs on an average of less than half a minute per nursery school morning

If time is at a premium the value of teachers' ratings should be considered In the present study one teacher's rating correlated 66 with all other records taken in the school and 58 with all the records taken at home and at school

Duration of crying behavior is a more valid index than number of crying incidents

For a cross section of the child's behavior in one situation such as a nursery school, somewhat the same picture is obtained regardless of the method used In a longitudinal study, comparison between incidence of behavior at different periods shows no such degree of relationship, because different factors modifying the child's behavior are in operation This is reflected in the low degree of consistency of behavior, which is interpreted as arising mainly from lack of consistency in the factors affecting the behavior

In conclusion, selection of method (or methods) should be based both on the characteristics of the behavior to be observed, and the record furnished of the context in which the behavior appears Without the latter, as will be shown in a subsequent report, sheer enumeration of incidents, or duration scores, are relatively meaningless

References

1 Arrington, R E Interrelations in the behavior of young children *Child Devel Monog*, 1932, No 8

2 Conrad, H S The California behavior inventory for nursery school children Univ Calif Syllabus Series, No 244

3 Jones, M C, & Burks, B S Personality development in childhood *Monog Soc Res Child Devel*, 1936, 1, No 4 Pp 205

4. Robinson, E W The measurement of children's emotions by various techniques M A Thesis, Univ Calif Library, 1933

Institute of Child Welfare
University of California
Berkeley, California

The Journal of Genetic Psychology, 1940, 57, 119-151

COMPREHENSION OF SPATIAL RELATIONS IN PRESCHOOL CHILDREN*

Clinic of Child Development, Yale University

EDITH MEYER[1]

A. INTRODUCTION

The investigation described here has been designed as a contribution to the genetic psychology of reasoning. It belongs to a series of studies of the evolution of notions in the child's mind. In his latest book, Piaget (11, 12) has described how notions like object, causality, space, and time are constructed progressively during the sensory-motor development of the child in the first years of life. These notions are closely connected in their evolution. As Poincaré (13) says, there is no such thing as space independent of objects. Relations between objects and movements of objects create space. Space is not the perception of a container but the perception of its contents. The notion of space can only be understood as a function of objects and all their relations. So the notion of space, if it is to be studied genetically, must be correlated with the comprehension of objects and of their properties.

The present study is designed to investigate, in certain selected situations, the evolution of the notion of space during the preschool years between 18 months and 5½ years.

There is a good deal of evidence to show that the preschool child is deficient in the comprehension of spatial relationships. Evidence can be found in studies of instrumental behavior such as have been

*Accepted for publication by Arnold Gesell of the Editorial Board, and received in the Editorial Office on June 25, 1939.

[1]This study was made while the writer held a Sterling Research Fellowship in the Graduate School, Yale University.

Acknowledgments are gratefully made to Dr. Arnold Gesell, Director of the Yale Clinic of Child Development, who kindly made available all the facilities of the Clinic and the Guidance Nursery and who read the manuscript, to Miss N. A. Hogan, Principal and to the teachers of the W.P.A. Nursery and the Kindergarten of Cedar Street School, New Haven, Connecticut, for their kind assistance in securing subjects and accommodations for the experiments, and to Mr. P. H. W. Hartmann, technician at the Clinic of Child Development, who helped in preparing the apparatus and drawing the charts.

reported by Gesell (3, 4), Charlotte Buhler (2) and others Other investigations of problem-solving behavior show a similar deficiency With techniques akin to those applied by Kohler (7) to chimpanzees, Alpert (1), Mathesen (8) and others have studied the general behavior of preschool children in problem situations André Rey (14), in his genetic studies of practical intelligence [some in collaboration with Lambercier (15)], has carefully analyzed the genetic progress of the child's structuring of the situations All these investigations show a growing ability of the child to adapt to various situations, to consider and foresee relations between objects, to effect appropriate changes in the position of objects and of himself, to analyze whole situations, to restructure them, etc In all such performances spatial relationships are involved.

B The Problem

The present series of experiments represents a preliminary approach to the whole problem The main points which seem to be worth investigating are the following What specific problems are involved for the child of preschool age in the comprehension of spatial relations? How do children of various ages respond to these problems? Is any genetic progress evident? What are the characteristics of such genetic progression, and what initial conclusions may be drawn with regard to the evolution of the notion of space during preschool years?

C General Method

The child was placed in various problem situations, which will be described in detail later Each situation included certain problems concerning spatial relationships The way in which a child handles the entire situation is liable to give us some information about his comprehension of the specific problem.

Behavior was observed carefully during the whole process of problem solving, since the purpose was not only to understand *what* the child does, but also to understand *why* he does it, and what underlying notions are revealed by his behavior. The whole process is of interest, trials and hesitations, reactions to suggestions, behavior in repeated and modified situations, all become important

D Subjects

Sixty-three children, between the ages of 18 months and 5½ years were examined (Table 1) Most of them were seen in a W P A

TABLE 1
Subjects

Age	18 mos	2 yrs	2½ yrs	3 yrs	3½ yrs	4 yrs	4½ yrs	5 yrs	5½ yrs	Total
Boys	—	5	1	4	2	4	4	6	4	30
Girls	3	4	5	7	4	5	2	1	2	33
	3	9	6	11	6	9	6	7	6	63

Nursery and in a municipal kindergarten, New Haven, Connecticut The children were mostly of Italian and Jewish origin They were all from the working class of New Haven The regular occupation of the father was variously reported to be common labor, truck driving, etc , most of the parents were on relief at the time of the investigation. The 18 month old children were seen at the Clinic of Child Development and had been referred for social reasons Some of the 2½ year old children, which were included, were pupils of the Guidance Nursery of the same clinic, these children belonged to a higher social economic class Although they were a little superior in some respects, their performance in our experiments did not differ to a significant degree from that of the children from classes with lower social economic background

E Experimental Findings

Although the study of the construction of the notion of space has to be made through the study of behavior, it must be kept in mind that the adaptations to spatial relationships as they are seen in behavior are not in themselves the center of the investigation They are only to be studied so far as they are apt to throw light upon the comprehension of these spatial relationships by the child It is the notion of space from the standpoint of the child and not from that of the observer which will be studied here

According to Piaget (12) two factors are essential in the construction of the notion of space (*a*) comprehension of the spatial relations between objects, (*b*) comprehension of the individual's

own shifts of position The experiments conducted in this study have been based on consideration of these two points

From among the different aspects under which the comprehension of the spatial relations between objects might be studied, two interrelated aspects have been selected for study here: (*a*) the fitting together of forms, (*b*) the comprehension of a moving object in relation to other objects

Comprehension of the individual's own shifts of position will be studied under the following two aspects. (*a*) coordination with a freely movable object, (*b*) coordination with an object whose movements are regulated

All these four aspects are closely related Some material bearing on each aspect could be found in the results obtained by every one of the experimental problems, but each group of results will be discussed primarily with reference to the one question to which they are most relevant

1 *Experiment I Comprehension of Spatial Relationships between Objects*

a The fitting together of forms

(1) *Problem 1 · Nested boxes* (Figure 1A)

(*a*) *Technique* The child is given a nested set of five triangular boxes (the largest of which is 6 by 6 inches, the smallest 2½ by 4½ inches) He is asked to take all the boxes out of each other and later to put them together again The way in which he manipulates the boxes is noted by the experimenter The same procedure is repeated This experiment was done with 58 children between 18 months and 5½ years

(*b*) *Results*. Behavior exhibited may be classified in three stages. The stages represent typical forms of behavior, which show a progressive evolution from entire disregard of forms and incomprehension of the impermeability of solid bodies to correct evaluation of form relations and a correct comprehension of graduation in forms and differences in size.

The first stage is most frequent from 18 months to 2½ years and disappears about three The following behavior is characteristic [2]

---

[2]*In the following descriptions of behavior the process of taking the boxes apart has been omitted, only the process of putting them together has been described*

In attempting to put the boxes together the child seizes two of the boxes closest to him and bangs and hammers them upon each other in any direction He does not adapt the forms to each other in any way It sometimes happens that he places a a small box, such as No 2, into a much larger one, such as

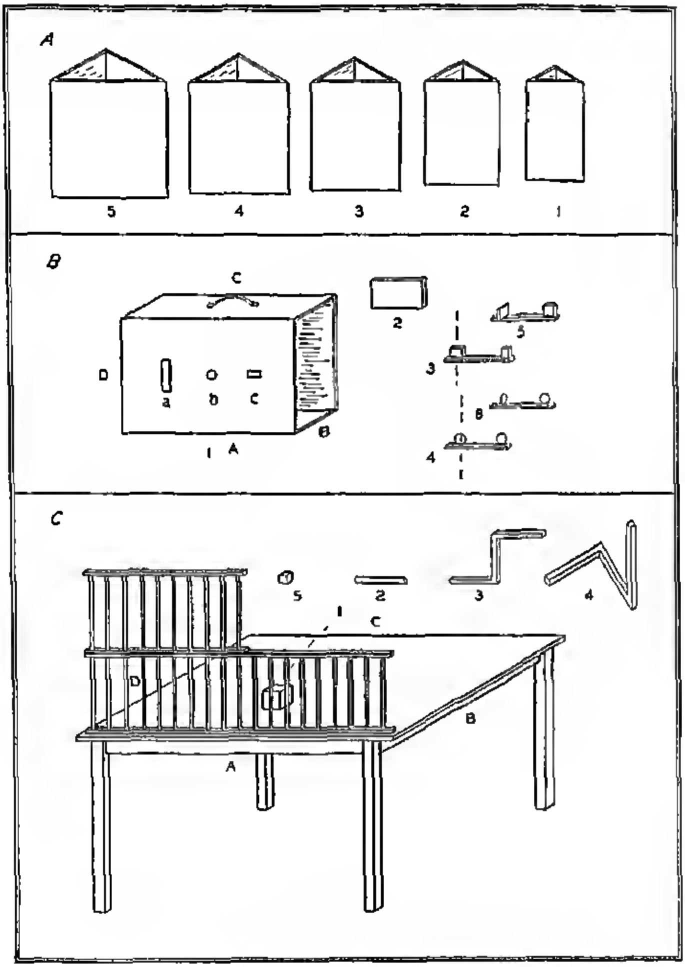

FIGURE 1

> No 5 He may then take another box, No 4, for example, and try to place it also in No 5 without realizing that No 2 is in the way He may continue such trials by repeatedly banging and hammering No 4 on No 2

Children in this stage show no appreciation of the forms and sizes of the objects they want to put together They do not perceive that one space only can be occupied by one object and that a larger object cannot be placed within a smaller one If they take one box out of another, in the course of reassembling the nest, they do it for the pleasure of the activity itself and not with the intention of removing an obstacle They choose the boxes to be handled quite without regard for their seriation

The *second stage* is most frequent at 3 and 3½ years It appears at 2½ years and disappears at about four It is characterized by a definite ability to learn from experience in the situation and to adjust accordingly (an ability which is lacking at the earlier stage) The behavior in this stage may be described as follows

> The child no longer merely bangs and hammers any two boxes against each other If he starts trying to put a larger box into a smaller one, he notices the impossibility and modifies his efforts He correctly holds two boxes in a position such that the axis of the one is a continuation of the axis of the other, in this position he presses the boxes together and turns them with a very characteristic rotating movement, until the smaller box reaches a position where it will click into the larger one If it happens that he has placed a small box, such as No 2 in a much larger box such as No 5, he tries only for a moment to place No 4 on top of No 2 and then immediately removes the obstacle, box No 2 In case he has already placed together Nos 2, 4, and 5 and wants to put No 3 in, he takes Nos 2 and 4 out to introduce No 3 and No 5 Or if Nos 2 and 3 are already out together, he works on Nos 4 and 5 separately, then when he is ready to place Nos 2 and 3 into No 4, he does not put them in together, but first separates them and may make some errors before getting them into the box properly

In this stage the child learns to adjust the forms to each other, he is not aware of all the relations beforehand, but he adjusts to them through experimentation. His prompt removal of an obstacle shows comprehension of the impermeability of solid objects. He applies the simplest principles of seriation in putting a smaller

object into a larger one, but he still cannot deal with a graduated series of forms, without making errors In his activity he does not take into account the relativity implied by a series The fact that a given box is large in relation to one size and small in relation to another size, is experienced by him but does not enter into his planning

The *third stage* is found in most of the 4-year-old children and is quite general from the age of 4½ on It involves correct adaptative behavior, with full comprehension of form-relations and of all the implications of seriation The typical behavior is the following

> Before bringing two boxes into contact with each other, the child chooses the right sizes and adjusts the position of the boxes so they will slide together easily He plans beforehand and does not only adjust empirically In case he has omitted, for example, Box 3, in placing the others together, he removes only the boxes which have to be removed from the next in order to introduce the omitted one He has no difficulty in starting his activity with Box 3, placing it inside No 4 and filling it with Nos 1 and 2 with a single movement

(2) *Problem 2 Performance box* (First part, see Figure 1B, 1 and 2)

(*a*) *Technique* In this experiment, the performance box devised by Gesell (3) and a rectangular wooden block were used. The performance box is a rectangular box (15½/10½/7½), which rests on one of the long sides and is open at the right end In the front side are cut three holes, a vertical rectangular hole (*a*) at the left, a round hole (*b*) in the middle, and a horizontal square hole (*c*) at the right The child sits at a table, and the box is placed before him with the open end facing him The rectangular wooden block is inside the box, lying on the bottom After the child has reached into the box and obtained the block, the box is turned 90 degree, so that the front containing the holes, now faces him The experimenter points to the hole and asks the child to pass the block through it If the child succeeds in turning the block in the right way, so that it enters the hole, he is asked to get the block out again and the test is repeated three times If he is not successful, the experimenter introduces the block into the hole and the child is asked to push it into the box, and then to get it out again The test was given to 60 children from the age of 1½ to 5½ The main problem again is the fitting of forms into each other

(*b*). *Results* With respect to the comprehension of the spatial relations involved, the behavior observed may be classified in three stages.

The *first stage* is found in all of the 18-month-old children and in some of those 2 0 old

> The child bangs the rectangular block flat against the front of the box. Repetition shows the same behavior The child does not know how to adjust the forms to each other.

The *second stage* is predominant at 2 0 and at 2½ but is also found at 3·0

> The child no longer bangs the block flat against the front of the box, but moves one end of the block towards the hole He may hold it in either a vertical or a horizontal position Often one corner of the block enters into the hole Sometimes the entire end enters into the hole by chance, but he does not exploit his success in repetition.

The *third stage* is found at the age of 3 0 and is quite general after the age of 3½

> The child shows full comprehension, he immediately introduces the rectangular block into the hole in the proper way, turning it in adaptation to the shape and position of the hole

(3) *Problem 3 Performance box* (Second part, see Figure 1B, 1, 3 4, 5, 6) In the next group of experiments, dealing with the fitting of forms into each other, a new aspect is introduced, the child's comprehension is studied by testing the flexibility of his activity patterns. Suggestions were made by the experimenter, and their effect was studied systematically. Also, situations were introduced which called for transfer, generalization, and differentiation of former experiences

(*a*) *Technique.* In the first experiment of the kind just described, the same performance-box, used in Problem 2 was employed In this problem a series of small instruments (3-6) are to be pushed through the horizontal square hole (*c*) One instrument, No. 3, is constructed of two small wooden blocks (square in shape) attached to the sides of a stick. The blocks turn freely on the axis, perpendicular to the stick When both blocks are turned parallel to the stick the instrument can be pushed through the square hole (*c*) No 4 is an instrument which is similar to No 3, except

that the square blocks are replaced by disc-shaped blocks No 5 is an instrument which is like No 3, except that one of the square blocks is so fastened, that it cannot be turned parallel to the stick; No 6 has disc-shaped blocks, like No 4, but one of them is again fastened Both Nos 5 and 6 are so constructed that neither can be passed through the hole.

Instrument 3 is presented to the child with the blocks turned so that the instrument will not fit into the hole The experimenter points to the hole (*c*) and asks the child to pass the instrument through it If the child does not succeed, the experimenter takes the instrument and slowly turns the blocks into the necessary position, making sure that the child observes all her movements, then she slowly introduces the instrument into the hole, pushes it in half-way and withdraws it After this demonstration the blocks are turned back to their original position and the child is asked to get the instrument in by himself The effect of the demonstration upon the child's activity is observed, and the same demonstration is repeated as many as three times if necessary After this, or earlier in the cases where no demonstration is needed, the instrument No. 4 is given without demonstration Later, several trials are given with No 5 and finally with No 6 These experiments were administered to 51 children from the age of 2 0 to 5½.

(*b*) *Results* Up to the age of four years correct introduction of No 3 into the hole (*c*) did not occur without demonstration. As will be shown later, even correct introduction is not complete proof of comprehension of the problem, which is how to turn the block in such a way that the instrument is of a form which can fit into the hole Only transfer to No 4 and rejection of Nos 5 and 6 after a short trial provide real proof With regard to comprehension of the spatial relations involved in the problem, the behavior may be classified into three stages

The *first stage* involved either of two forms of behavior which differ slightly

> *a* (Predominant in the children of 2 0 and 2½). The child tries to introduce the instrument No 3 in an entirely unadaptive way Often he holds the instrument parallel to the box instead of perpendicular to it Even if he holds it in the proper position, he only works on one of the blocks and is not able to adjust even it properly Demonstration has no effect and seems

to be considered only as an undesired interruption No 4 is handled in the same unadaptive way Nos 5 and 6 are tried for as long and with as much effort as the other instruments

b (Predominant in the children of 3 0 and 3½) Here the child shows the same initial behavior, but differs in his reaction to the demonstration He immediately imitates the experimenter by turning the blocks After doing this for a while, he tries to introduce the instrument into the hole without any concern for the position in which his turning has left the blocks The turning activity seems to be understood as a sort of ritual which is expected to make the introduction of the instrument more likely. The same turning activity, without concern for its real effect, is transferred to No 4 Later, when the child notices that one block on No 5 and No 6 cannot be turned, he tries to introduce these instruments into the hole just the same

All the children belonging to this first stage show an incomprehension of the spatial relations involved in the problem These included under "*a*" are not influenced by any demonstration, and no transfer takes place, those included under "*b*" show a certain aptitude to follow a demonstration, but they only imitate the demonstrated movement and transfer it to other situations without comprehension of its function

The *second stage* is predominant in the subjects of 4·0 and 4½ It reveals a more adaptive empirical handling, leading to final success but not to entire comprehension The typical behavior can be described as follows:

The child shows an experimental attitude of being ready to profit from immediate experience. He pushes the instrument towards the hole and tries to get it in, making various trials and patiently turning the blocks Most of the children succeed after long efforts Repetitions of the problem result in the same type of response rather than an immediate solution A demonstration has to be given to only a few children When given, a demonstration does not seem to suggest the entire solution to the child but only a conviction, that the problem can be solved by experimentation The child transfers the same attitude to No 4 and manipulates No 4 until he gets it in Nos 5 and 6 are tried, but the child's enthusiasm for them diminishes, when he discovers that one part of these instruments is rigid and does not yield to his efforts at adjustment But he does not entirely reject these instruments If asked if he expects to be success-

ful, he is likely to declare that he cannot know as yet and must try further

In this stage of an experimental attitude without entire comprehension, some progress in spatial comprehension may be observed The blocks or discs are correctly adapted to the hole, but the interdependence of the two blocks or discs through the stick, to which they are attached, is not yet understood.

A *third stage* is found in the children 5 0 and 5½ years old Progress in comprehension is shown in the following behavior.

> After the first trial, most children immediately adjust the blocks on No 3 in the right position before again bringing the instrument up to the hole In the cases in which a demonstration is given, it immediately suggests the correct adjustment of the instrument No 3, before any new efforts are made to get it into the hole When No 4 is presented, it is immediately studied and then adjusted in the same anticipatory way Nos 5 and 6 are rejected after the discovery that "it cannot be turned," that "it is too big," etc (Some children in this group show a specific attitude with instrument No 5 which might be mistaken for lack of comprehension After having discovered that the ordinary procedure of turning the parts parallel to the stick does not work, they try hard to get the instrument in anyway, apparently thinking, that the task can be done by another trick, which they are supposed to discover Questioning readily reveals this so-called "trick-attitude," which is generally found only with the oldest children) The real comprehension characterizing this stage, is evident in the foresight demonstrated in the handling of the objects, in the transfer to No 4 and in the rejection of Nos 5 and 6

(4) *Problem 4 Objects behind fence* (see Figure 1C) Another series of experiments was arranged for the study of suggestions, transfer and differentiation

(*a*) *Technique* The child is placed in front of a large table at side *A* Interposed along the edge of the table at side *A* rises a fence constructed of thin wooden posts, set at intervals of three inches The right half of the fence is 10 inches high and the left half is 20 inches high Sides *B* and *D* of the table are open and accessible, whereas side *C* is placed against the wall A big red cube (see Figure 1C, 1), too large to be passed between the posts is

placed behind the low part of the fence and the child is asked to get it. The possible solutions are to lift the cube over the fence or to push or lead the cube along the table behind the fence to one of the open sides, *B* and *D*, and obtain it there If the child does not reach either of these solutions, suggestions are introduced The first suggestion consists of the experimenter's lifting the cube over the fence and handing it to the child, the cube is then replaced in its original position on the table

When the child has reached a solution, a second problem is posed, by placing the cube behind the higher part of the fence. If no solution is reached by the child, the following suggestion is given the cube is moved over to side *B* by the experimenter, given to the child, and then replaced in its former position Later various other objects are placed behind the fence—a long rod (see Figure 1C, 2), two different angleblocks (*C*, 3 and 4) which can be passed between the posts, if turned in the right way and in some cases a small cube (*C* 5) which can be passed between the posts These experiments were done with 50 children from 18 months to 5½ years

(*b*). *Results* The results will be discussed here with regard to the spatial relations between the objects involved (objects and fence), later the results of this same experiment will be discussed with respect to the shifts of position involved Three stages may be distinguished which mark a growing comprehension of the relations between objects.

The *first stage* is found from the age of 18 months up to 2½ years It simply involves efforts to obtain the objects by direct prehension

> The child drags at the cube only at the spot where it is already located, and tries to pull it between the posts He does not seem to understand that the cube is too large to go through. Suggestions have no effect, the child is happy to get the cube, when the experimenter gives it to him, but he is not at all interested in the trajectory made by the cube in the hands of the experimenter The other objects are dragged and forced through the fence, with no attempt to turn them adaptively

There is no comprehension of the spatial relations involved between the objects, only the child-object relationship is active and dominant

The *second stage* is found at the ages of three and four years Its characteristics are manifested in varying forms of behavior

> Some of the children, after having tried to secure the object by direct reaching, arrive by themselves at the solution of lifting the cube over the low part of the fence, others arrive at this solution after a demonstration. In the later problems, this pattern of lifting objects over the fence remains rigid and undifferentiated.
>
> (*a*) Some children, when the cube is placed behind the high part of the fence try to lift the cube over, even though the fence is here too high to permit their doing so; they do not find any new solution for this situation. New suggestions have no effect.
>
> (*b*) Other children move the cube over to the exact spot behind the low part of the fence, where they have lifted it over before. They do not see the possibility of moving the cube to the open side *D* which would be the most direct solution. Suggestions in this direction usually have no effect.
>
> (*c*) Another form of rigid behavior is manifested in the following way: with new objects, the same pattern remains unmodified. The long rod and the angle-blocks are lifted over the low part of the fence, even though they could easily be pulled between the posts. In many cases, even the small cube was lifted over the fence.[3]

The characteristics of these forms of behavior seem to be a rudimentary comprehension of spatial relations between objects. The child finds one solution either by himself or as a result of suggestion.[4] His behavior is characterized by a rigidity of the method of solution. The one successful pattern adopted during his activity is applied to all situations which are somewhat similar, and he does not differentiate his activity pattern according to the special spatial constellations into which objects are placed in the several situations.

The *third stage* is found in most of the children at the age of 4;0 and later. It involves full comprehension and ability to differentiate the activity pattern in accordance with the situation.

---

[3]Here it should be mentioned that the possibility may be excluded that the children did not understand that they were allowed to use any kind of solution. The whole behavior of the children showed how eager they were to get the objects, and that they were not at all handicapped by inhibitions of this kind.

[4]The results confirm the findings of A. Rey (14) that suggestions are effective only when the child is very near to that point in his mental development at which he would reach the same solution spontaneously. At earlier stages, the same suggestion has no effect, because the child is not mature enough to respond to it.

> Both types of solution (lifting the object over the fence and pushing it to a free side of the table) occur spontaneously. When the solution which he has not discovered for himself is suggested to the child, he adopts it immediately. The mode of solution is very definitely adapted to the variations in conditions. The child seems to plan the best method for each object and each situation. He lifts the large cube over the low part of the fence, but when it is behind the high part of the fence, he pushes it to the nearest open side of the table. The rod and angle-blocks are turned adaptively and pulled through between the posts.

b. The comprehension of a moving object in relation to other objects. Comprehension of space as well as of time[5] grows out of the comprehension of movement. Comprehension of space is arrived at when shifts of position are understood. In this section, comprehension of the movements of objects will be discussed; later, the child's own shifts of position, and their coordination with those of objects, will be investigated.

At the beginning of his sensory-motor development, the child does not understand that a moving object follows a certain path in space. As development proceeds, the child seems to understand the relations of movement to the surrounding space. Piaget (12) describes a progressive adaptation to rectilinear movements. In the first months of life, the child gains the ability to follow slow rectilinear movements.[6] Around seven months he is also able to refixate an object which is moving so rapidly that he momentarily lost his fixation (12, pp. 22-25). At this same age, he is still not able to maintain the proper sense of location for an object which disappears in a slow movement behind a screen and which will reappear at a point which could be predicted from anticipation of the path it is following; the child reverts to a fixation of the point at which the object first appeared (12, p. 120). Piaget concludes that in the case of a rapid movement it is only the movement and not the shift of position as such to which the child accommodates (12, p. 25). Shifts of spatial positions are not yet understood; the child continues to look for a disappearing object at the wrong place until the age of 9 and 10 months.

---

[5]Study of the evolution of the notion of time is reserved for a later paper.
[6]Gesell (4) lists such behavior at 6 weeks.

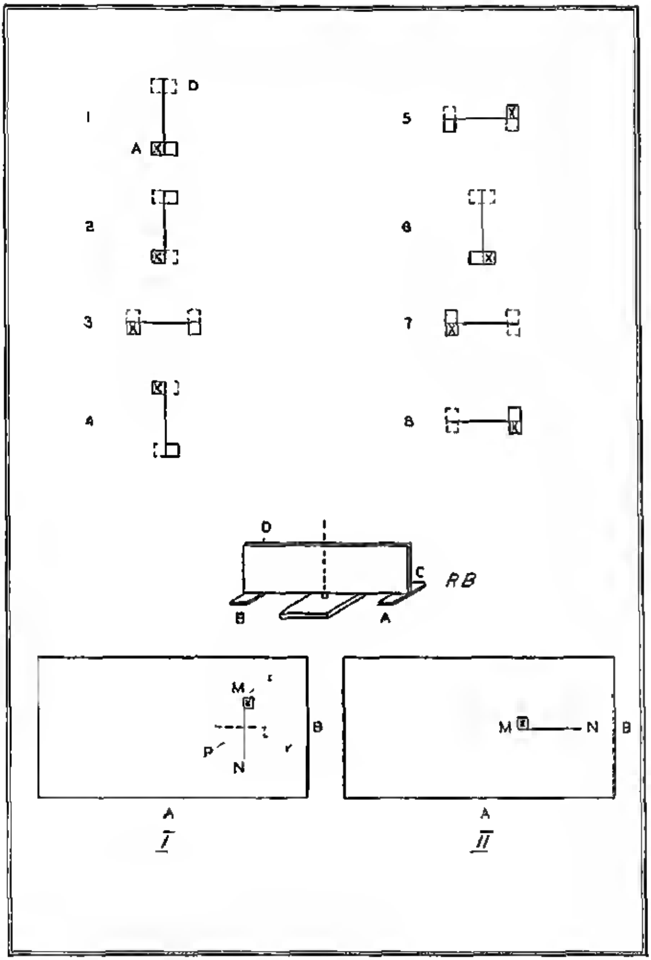

FIGURE 2

A similar evolution has been found in the present investigation to obtain at higher levels with the comprehension of displacements from rotation. If, for example, a rotating lever is shown to the child, he understands the movement at an early age, only as a dynamic and unspatialized movement independent of space It is only

at a later stage that the child realizes that the lever passes definite spatial points in the course of its movement This evolution has been studied in the following experiments

(1). *Problem 5 Rotating board* (see Figure 2RB)

(*a*) *Technique* A board six inches high and 14½ inches long is mounted on a pivot. At each end of the board are attached two small platforms (*A, B, C, D,—A, B* on the one side, *C, D* on the other side—see Figure 2RB) This apparatus is placed on a table in front of the child. A colored chip is placed on one of the platforms while the child watches, and it is covered with a cardboard A similar cardboard is placed over another platform which is empty The child is asked "to watch where the button goes" Then the board is slowly turned 180 degrees The child is asked to get the chip. The first two or three trials are considered as preliminary. Then follows a series of variations of the same experiment, in each variation the board is turned 180 degrees in the same clockwise direction, but the other conditions are changed in the manner shown in the diagram (see Figure 2, 1 to 8). The whole series is repeated, when any of the solutions have not been correct After this series the child is asked to make predictions "Where do you think that the chip (button) will go when I turn?" After the child has guessed, the experiment is executed Sufficient motivation was provided in all these experiments by the wish of the child to get as many chips as possible. (He was allowed to keep the chip each time that he found it without making an error) With younger children the pleasure in finding the chip provided enough motivation in itself to keep their interest aroused for a relatively long period

The experiments were done with 45 subjects from 2 0 to 5½ years It was found that prediction could only be asked of children who were at least four years old.

(*b*). *Results* The several variations in the experiments allow a comparison of the corresponding results. Sometimes negative results in one variation permit an understanding of positive solutions of another The analysis of the responses of the children in these experiments has led to a definition of four stages in the evolution of the comprehension of displacement from rotation

The *first stage* is predominant at the age of 18 months and 2 0 years and is still found at 2½. It lacks entirely any representation of the shift of position which has occurred. This is manifested in several different forms of behavior.

a Some children look attentively at the moving board, but when it stops they look for the chip in quite absurd places They look under the apparatus, at the pivot, under the rubber suction cups on which the base of the apparatus is mounted[7] At the moment when they lose sight of the chip, its movement becomes incomprehensible to them

b Other children follow the chip until it has disappeared and then look back immediately at the place, in relation to themselves, that it occupied before the movement started When, for example, in Variation 1 (see Figure 2) the chip has originally been placed on Platform *A* and the apparatus is rotated 180 degrees, they now look at Platform *D* which is at the position formerly occupied by *A*

c Other children, once the board has stopped, try to turn it back again They are forbidden to do so, and they do not find any other solution They understand only that the chip has disappeared because of the movement of the board and act as if the only way to get the chip back is for them to undo the movement

The general characteristic of this stage is a definite lack of representation of the shifts of position involved. The chip disappears, but its new location cannot be deduced from observation of the movement and of the resulting position of the board No relation exists for the child between movement and change of position, movement is understood as a movement but not as a shift of the spatial position of an object

The *second stage* is found at 2½ to 3½. It is manifested in the following types of behavior.

a Many children follow the movement until the chip disappears and then wait eagerly for the first covered platform to reappear as the movement continues As soon as any cardboard comes into sight, they expect the chip to be under it. So in some variations (as in 4 and 5, see Figure 2) their choice is the correct one, while in others (such as 7 and 8) their choice is incorrect They know that the desired platform must be expected at a different point from before; but they only take into account the one desired platform and neglect the relations of the whole situation They do not take into account the empty covered platform and its changed position

[7] In another experiment, the suction cups are used to fasten the board to the table top

b Other children correctly look at the rear of the board, when the turning has left the chip there But they do not look in a definite position, either of the two platforms at the back seems right to them In Variation 7, for example, they are likely to choose *C* as well as *D* Many of the correct solutions thus do not involve complete comprehension, this is evident from remarks made by the children (such as "it is way back," etc), as well as from comparison with incorrect solutions

In both forms of behavior the characteristic of the stage may be seen in the beginning of a comprehension that the chip has changed its position through movement and that it has to be expected at a new place. This new position is anticipated only in a rather undifferentiated way In the group described under "*a*" the chip is expected simply somewhere along the path it has begun to take; in the group described under "*b*" a shift from the front side to the back side of the board is understood The progressive shifts of spatial position made by a rotating object are not taken into account The child still lacks correct representation of the invisible portion of a moving object's path In a transitional stage the following behavior is found

a The child follows closely with his eyes or even with his finger the movement of the platform which holds the chip He continues to follow its movement after it has disappeared and he resecures the chip at the right place So he seems to understand even the invisible movement of the object But if in any one trial he does not attend to the movement quite as closely, he is incapable of finding the chip again when the movement has stopped The new position of the board does not indicate to him where the chip must have stopped, his behavior reverts to one of the forms which have been described for the second stage

The *third stage* around the age of 4 0 shows decisive progress in comparison with the previous stages, it reveals a true comprehension of the displacements from rotation.

In his behavior in the experiments the child does not show any difficulties when he has to resecure the chip after the board has turned He does not need to follow the movement closely with his eyes or fingers, it suffices for him to see the initial position, the beginning of the movement, and the final position,

He is able to reconstruct the new position of the chip from the new position of the board, he is able to represent the spatial components even of movements which cannot be seen, thus he shows that he has understood the spatial relations involved in the movement

But the interesting feature of this stage is that the representations are limited in their function Most of the children up to the age of 5½, though able to resecure the object when the movement has been completed, are still incapable of predicting the future position of the chip before the board has yet been turned the usual 180 degrees. Some children make the same mistakes as the children in the previous stages in situations with completed action, they point to some indefinite place on the table outside the radius of the board, or some place in the vicinity of the board but bearing no relation to the platforms (i e, the middle of the board, the rubber suction cups, etc) Other children assume correctly that the future position of the chip will be on one of the platforms of the board, but point indiscriminately to either one of the platforms which are on the correct side of the board [8]

In general, the children at earlier stages are incapable of predicting and do not even understand the question Each activity and event seems new and unique to them and happens without predictable relations or laws But at the present stage the question is accepted and the child agrees that a certain new position can be expected, so long as the usual practice of turning the board 180 degrees is continued But his representations fail to function when they must be detached from actual concrete phenomena His representation is adequate when he is required only to reconstruct a past condition from the presently given state of affairs It is only at a fourth stage that correct predictions and entire comprehension are found with correct representation of all the position changes involved, even

[8]André Rey (14, p 162) has found similar errors of prediction in a comparable problem which had been devised to study the comprehension of simple mechanical problems in children His apparatus consisted of a ruler, fastened in the center and free to rotate, at one end of which was attached a string and at the other end an object The child was asked to foresee the position of the object, if one pulls at the string His subjects gave a variety of predictions similar to those described here They gave 100 per cent wrong predictions between the age of 3½ and 4½, 75 per cent wrong predictions up to the age of 5½, and still 27 5 per cent errors between 5½ and 6½

without any relation to a completed action. This stage is only in a few cases found within the age-levels studied here.

2. *Experiment II: Comprehension of the Individual's Own Shifts of Position*

a. Coordination with a freely movable object. The next problem is to investigate the way in which the child comes to understand his own shifts of position. Piaget (12) points to the fact that at the early sensory-motor stages of mental development this comprehension does not exist. He says that at early stages the child begins to discover spatial relationships among objects, and considers himself, if not outside space, nevertheless as an invariable center. He is not as yet an object which is like other objects and whose shifts of position take place in relation to the shifts of position of other objects.

In some informal experiments with children from 18 months to 2½ years it has been possible to add some examples to the work done by Piaget (12). Results show a growing ability of the child to make detours which require representation of his own change of position as relative to objects in space.

(1) *Problem 6. Bell behind performance box* (see Figure 1B).

(*a*). *Technique.* The child stands at *A* in front of a small table on which the performance box (*B* 1) is placed. The experimenter sits at the left side of the table, so close to it that there is no space left between her and the table. She shows an object (a bell) to the child, and then moves it slowly around the left side of the box and hides it in back of the box at the point marked *C*. The child tries to get the bell back.

(*b*). *Results.* A marked tendency was observed in all subjects 18 months old, and in many of the 2:0 year old to try to obtain the object by following the same path which it has taken, namely, from *A* through *D* to *C*. On discovering that they could not follow this path the children of 18 months gave up all attempts to get the bell, started to fuss, and did not see any other solution. Most of the 2-year-olds, after some hesitation, discovered that they could go around the other side (*B*) of the box to get the bell. The 2½-year-old group immediately perceived the situation and realized that they could not follow the object around the left of the box but could obtain it by going around the right side *B*.

(2). *Problem 7 Getting block out of performance box*

(*a*) *Technique* It will be remembered that in the procedure of Problem 2 the child obtained a rectangular block by reaching into the open end of the performance box The box then was turned 90 degrees, so that the side with the holes was now in front of the child, the child pushed the rectangular block through one of the holes and then he was asked to get the block out again

(*b*) *Results* All the 18-months-old children as well as the 2 0-year-old children try at first to get the block back through the hole through which it has disappeared The 18-months-old children persist in poking in the hole and do not see any other solution, but most of the 2·0-year-old children finally discover that they have to move over to the open end *B* of the box Some of them need one suggestion by the experimenter before they can find the solution themselves

At early stages the child only learns gradually to find an object by following another route than the one which he has seen the object take In other words, he only learns gradually that in a square *ABCD*, whose center is blocked by the box, he can go from *A* to *C* by the path *ABC* as well as by the path *ADC*. In such a detour the child abandons the object-subject line which has been established by his experience and yet understands how to get hold of the subject. He not only represents the geometrical arrangement of the spatial points but also his own changes of position as movements between such points

(3) *Problem 8 Objects behind fence* (see Figure 2C)

(*a*) *Technique* (See page 129)

(*b*) *Results* The results of Problem 4 (objects behind fence) have until now only been discussed in the light of spatial relations between objects and their forms Now the same data will be looked at from the point of view of position changes, especially those of the self

In comparison with the problems just described the "object behind fence" experiment includes a new aspect so far as shifts of position are concerned Whereas in Problems 6 and 7 the child has to adapt his own movements to a certain definite position of the object, he now has to effect changes in the position of the objects as well as of himself. His own shifts of position have to be included in a system with other shifts

In the *first stage* (around 18 months to 2½ years) which has been described above the child tries to pull the cube through the gate directly at the place where he finds it. He does not consider any change in position either of the object or himself beyond his direct reaching for the object

In the *second stage* at 3 0 to 4·0 years, preference was found for lifting the cube over the top of the fence There a detour is made with the object, but the subject maintains his own original position

In the *third stage* (at the age of 4 0 and later) the child makes a detour himself by going to one of the open sides of the table (*B* or *D*) He not only moves the object but he also takes into account possible changes in his own position A spatial point *B* or *D* outside of the line of force from object to subject becomes the center of the field, both subject and object are directed towards this point and are supposed to meet there

This achievement is a very decisive one for the comprehension of spatial relationships The child ceases to be a privileged center and begins to consider himself with reference to the objects

b Coordination with an object whose movements are regulated

(1) *Problem 9 Rotating board on table* (see Figure 2RB and I and II) The last problem which will be studied is that of how the child comes gradually to coordinate his own positional changes with those of objects whose movements are regulated These experiments were primarily inspired by occasional observations

Preschool children of early ages commonly run behind a toy-train when they want to catch it They do not foresee that they would catch it more easily by crossing rather than following the path of the object This solution requires representation of the shifts of position of the moving object and of the self The coordination of both seems to be important in the evolution of the notion of space, and some experiments have been devised for its study.

(*a*) *Technique* The board (*RB*) described under the technique of Problem 5 is placed on a large table The base on which the board pivots is now attached firmly to the table by rubber suction cups The children cannot drag it along the table at all and the only movement possible is a rotary one The distance of the board from the edge of the table is adjusted at the beginning of each experiment, according to the size of the child.

The first situation (I) is as follows the distance of the board

from side *A* of the table is such that the child is just able to reach the closest tip of the board. The distance from the right side *B* is such that the board can only be reached after it has been turned 90 degrees A chip is placed on platform *D* in *M*, the child is standing at *A* and is asked to get the chip After a brief moment (in which most of the children try to reach directly) the experimenter demonstrates, by moving the end *N* once back and forth, that the board can be rotated The instruction is "you get the chip, you can do whatever you like to get it"

In a second situation (II) *N* is turned towards *B*, the chip is placed in *D* The child is at *A* This situation is usually repeated several times.

The first situation (I) is repeated with the child in *B* and the second situation (II) with the child in *B* Sometimes the child needs encouragement at the first trial to "go where he likes"

The experiments have been given to 50 children from the age of 2½ to 5½ years old

(*b*) *Results* In the results three different groups of solutions may be distinguished as stages in a progressive evolution towards comprehension of the continuity of the position changes involved in a rotating movement and towards facile coordination of those changes with the shifts of position of the self.

The *first stage* is represented by a group of responses given by all children of 2½, most of the 3 0 years old and 50 per cent of the 3½ years old

In this group of responses characteristics are found which are already familiar from the description of behavior on other problems at early stages

> There is a direct object-subject line and incomprehension of the displacements from rotation The child does not understand that the rotating of the board can bring the desired object into his reach He continues to reach directly towards the chip He shakes and pulls at the nearest end of the board trying thus to drag the chip in a direct line towards himself This is impossible, because of the attachment of the board to the table Sometimes this tugging leads, by accident, to a rotary movement which may bring the chip almost within the child's reach, if this happens he immediately gives up the pulling and reaches directly towards the chip even though he still cannot reach it He does not see that the rotary movement of the board was the

reason for the increased proximity of the chip, and that he should continue it When the pivoting of the board is demonstrated, the child may imitate the movement of the experimenter, but it is evident that he does not see any relation between the pivoting movement and the getting of the chip

The *second stage* is represented by a group of response characteristic of the children of 4:0 years The traits already revealed in other problems appear again here The movements effected are directed towards the subject as privileged center and not towards an independent point in space The rotating movement is still only very roughly spatialized but is understood as a method of bringing the chip within reach. The following behavior is characteristic.

The child imitates the experimenter in turning the board back and forth, he often experiments a short while with the board and seems to discover that the chip disappears when the board is turned to the right, and remains in view when turned to the left. Then he tries to effect an entire 180 degrees turn to the left This seems to be for him the only effective method of bringing the chip within his reach He does not try to stop the board when the chip points to *B*, from where he could then easily reach it, he remains at *A* and tries to get the chip near that point even though the reach is at best more difficult than from *B* When he has not pushed the board vigorously enough, it stops in a midposition, that is with *M* pointing to *B*. He tries to reach from *A* towards the chip only for a moment and finding that he cannot reach it, tries to push *N* further, even when this end is far away from him Mostly he does not even see that he could reach the chip at *B* Only sometimes when he has already entirely given up all attempts, he goes around to *B* as if guided by a new perception, and gets the chip at *B*, but it is very obvious that this is no satisfactory solution for him It is only a way out, for at the repetition he tries to achieve the entire 180 degrees turn as before Situation II is solved in a similar way to situation I The child goes from *A* to *B* and tries to turn the board 180 degrees in the direction which keeps the chip in his view He tries hard to push *N* further when the board stops in the midposition with *N* pointing towards *A* In a few cases, the child after many such trials finally goes over to *A* and tries again to make this difficult 180 degrees turn He again does not stop the board at *B*

The *third stage* may be distinguished in a group of responses given

by most of the 4½ year old children and almost all of 5.0 and 5½. Again we find here traits noted in the presence of similar problems. The progress of the stage lies in increased coordination of the movements of the self with those of the objects, and comprehension of the progressive and continuous position changes involved in the rotating movement. The characteristics of the stage become especially evident in the behavior in Situation II.

> The child goes from *A* to *B*, turns the end *N* 90 degrees towards *A*, goes back to *A* and guides the board slowly and carefully till the end *M*, bearing the chip, approaches *B*, then he returns to *B* and receives the chip there. The child chooses that direction of rotation in which the chip is always visible. He does not make efforts at *B* to turn the board 180 degrees, because he foresees that the board will pass some spatial points which can be reached. If he does not foresee these specific points themselves, his attitudes show that he tries to be alert to them.

Here is, however, a primitive residue in the fact that the child turns the board only in the direction in which the chip remains visible to him; this forces him to make many more movements then are necessary for the perfect solution of the problem.

Theoretically it could be assumed that at a higher stage the child would make only the essential movements. He then would immediately direct the end *M* towards *A*, turning the end *N* 90 degrees towards the right. In this case a representation of a partly invisible movement would be projected beforehand and coordinated with representation of self movements, and the child would no longer hesitate to see the chip momentarily disappear. But only a very few such instances have been found in age-ranges studied here. Logically it may be predicted that this more economical solution would occur at a higher level. However, the experimental set-up used here may be not adequate to elicit such a simple solution with older children if the motivation for the most economic solution were not strong enough with them.

F. Discussion

An analytic investigation has here been made of the development of the notion of space in the preschool child. Children were presented with various problems, the solution of which seem to represent

TABLE 2

Experimental situations	Problem 1 Nested boxes (see Fig 1 A)	Problem 2 Performance box (1 Part) (see Fig 1 B, 1, 2)	Problem 3 Performance box (2 Part) (see Fig 1 B, 3, 4, 5, 6)
First stage (Pratical space)	No appreciation of forms and sizes of objects (no comprehension for the impermeability of a solid) No regard for seriation 18 months—2½ years	No appreciation of forms 18 months—2 years	No appreciation of forms (demonstration without effect) 2—2½ years (demonstration elicits imitation without comprehension, transfer of imitated movements to other situations) 3—3½ years
Second stage (Subjective or empirical space)	Adjustment to form relations through experience Comprehension of simplest principles of seriation (no comprehension for relativity implied in seriation) 3—3½ years	Beginning of adjustment to form relations 2—2½ years	Partial adjustment to form relations (Empirical attitude, no planning beforehand) 4—4½ years
Third stage (Objective space)	Correct comprehension and adjustment to form relations Comprehension of all necessary implications of seriation 4 years	Correct comprehension and adjustment to form relations 3½ years	Correct comprehension (demonstration has full effect, transfer and differentiation with regard to other situations) 5—5½ years

TABLE 2 (*continued*)

Experimental situations	Problem 4 Objects behind fence (see Fig 1 C, 1, 2, 3, 4, 5)	Problem 5 Rotating board (see Fig 2 RB, 1-8)	Problem 9 Rotating board on table (see Fig 2 RB, I, II)
First stage (Practical space)	No comprehension of relations between objects		
		Rotating movement understood as dynamic movement, not as a change of position	No comprehension of changes of position involved in rotating movement
	No detour (direct prehension) 18 months—2⅓ years	18 months—2½ years	No detour (direct reaching) 2½ years—3½ years
Second stage (Subjective or empirical space)	Rudimentary appreciation of relations between objects (rigidity of the method of solution, generalization but no differentiation of the method of solution)		
		Rough evaluation of changes of position from rotation (no representation of invisible portions of a moving object's path)	Rough evaluation of changes of position from rotation
	Detour of object (subject maintains its position as immovable center) 3—4 years	2½—3½ years	All movements effected are centered on an immovable subject 4 years
Third stage (Objective space)	Comprehension of relations between objects (differentiation of method of solution in accordance with the situation)		
		Comprehension of changes of position from complete rotation (no correct foresight) 4—5½ years	Comprehension of continuous changes of position involved in rotating movement
	Changes of position of object and subject are coördinated 4 years		Coördination of changes of positions of subject and object 4½ years

essential elements in the construction of the notion of space. The children's comprehension of these problems was studied by qualitative analysis of their behavior. The results are summarized in Table 2.

Comparison of children of all ages made it possible to distinguish between features of behavior which merely reveal individual differences likely to appear at all levels, and other features which seem to be characteristic of a certain level of mental maturity. In accordance with the purpose of this study, genetic trends have been abstracted out of the concrete behavior, and the individual differences, which in themselves pose significant problems for research, have been deliberately set aside The individual differences do not appear to overshadow or invalidate the genetic trends In the following paragraphs the genetic findings will be reviewed.

From these facts it is possible roughly to sketch the notion of space which underlies the behavior in the described situations which call for representation and analysis of complex spatial relations

Up to the age of 2½ the child behaves in these situations as if he moved in what may be called with Piaget a "practical space," in which objects are only present to satisfy his activity All his needs are centered around the exercise of the activity patterns which he has so recently acquired during the sensory-motor development of the first two years. Serving as material for the exercise of a self-satisfying activity is the only function of the objects around him The objects mainly have reality in relation to the child himself, who neglects relations between objects This has been seen to hold true of form-relations, as well as for all the spatial relations involved in movements other than rectilinear which take place independently of the child.

The ages between 3.0 and 4 0 are those in which he moves in a "subjective" or "empirical" space A definite advance is found in comparison with the former stage. The child begins to show an interest in the objects themselves, though he is still centered upon himself and upon his acting upon things. But whereas his activity hitherto forced its pattern rigidly upon things, at this stage his activity adapts its patterns to the conditions of the environment Objects no longer constantly block and resist the flow of the activity, but they are integrated into it by gradual modifications of the primitive patterns The child is able to profit from experience and

to transfer a pattern, just acquired, to new situations. Though he knows how to generalize his experiences, he still has difficulty in differentiating and reorganizing his patterns according to the special conditions of the objects. At this stage he adjusts to the environment through its pressure on his activity. He still feels himself to be the center of this environment. He moves an object only towards himself, when an object is rotated away from him, he does not consider that the object has an independent path of its own which includes more than the portions which are visible to him. In space at this age, object relations are only roughly evaluated and have reality only so far as they are connected with a subject who considers himself as a sort of privileged center of the universe and not as a part of it.

After the age of about 4 0 the child starts to move in an "objective" space. Interesting changes take place in the relations of the child to his environment. Activity loses its predominant position and the objects and their conditions replace it. The child sets out to live in this world of objects and no longer to make things exist merely for his own satisfaction. His patterns of activity are adapted to the objects. They are constantly differentiated and reorganized in view of the properties of the objects, which he recognizes even before they force their pressure upon him. He is no longer the center of his world but a part of it. He understands the spatial relations within this world of objects, he understands their independent movements and he knows how to coördinate his movements with theirs. He now acts among things and no longer only upon them. Space at this age is an objective space, whose relations are understood apart from all relations to the child himself. He knows that he has to consider himself as one object among other objects in space, and he understands the relativity of the positions of the objects.

The evolution described here in some of its representative steps is a progressive one. The borderlines between the stages are fluid. Transitory periods are intermediate between the different stages. Each stage carries with it the residues of a foregoing stage and the seeds of the one to follow. The age values which have been given here are meant to mark nothing else than the age which is the most representative for the stage. As may be seen in Table 2 a stage may be reached earlier with one problem than with another, according to the difficulty of the problem.

The process of progressive objectification of the notion of space in the child's mind goes along with the progressive shift from an egocentric point of view to a relativistic one. It is not the first time and not the last one during his mental evolution, that the child goes through this development from a practical, through a subjective, to an objective notion of space

As Piaget has pointed out, the child has already once gone through this evolution during his first two years of sensory-motor development At the earliest stages of this development, the newborn perceives no permanent object. Only pictures surround him, they are related only to the different aspects of his primary activity (vision, audition, prehension, etc.) and have no reality outside of it At an early stage there seem to be as many practical spaces as practical activities The patterns of primary activity are coordinated with each other during the first months of life, and along with this coordination goes the coordination of the different practical spaces to a common space. In a subsequent stage of subjective space, the child comes progressively to differentiate between his movements and those of the objects, is able to do this so long as the movements of the objects are in immediate relation to his own activity He comes also to notice simple relations of objects such as behind, in front of, upon, etc About the end of the first two years he begins to represent some of his own movements as belonging to a system of spatial relations which includes both his own movements and those of objects In this way, the stage of "objective" space is reached on the restricted field of *sensory-motor* activity

A previous study by the present author (9) has shown similar stages in a problem concerned with spatial relations on the level of *reflective* thought. In that problem the child has to consider that his own point of view is only one among other points of view and that a person with another spatial point of view sees things differently A stage of "practical" space was found below the age of 7.0, in which the child does not understand that a scene depends upon the point of view In a following stage of "subjective" space (7 to 8 years) he comes to distinguish some of the relations between the objects and knows that from another point of view the scene is different, without understanding which relations change Through transitional stages he reaches the stage of objective space (9 to 10 years) in which he understands both the relations between objects

and the relations between the different points of view and the corresponding appearance of the whole scene

Comprehension of spatial relations occurs on different levels of mental evolution, according to the difficulties of the problem On whatever level it appears, it progressively develops by a continuous genetic process of which the developmental stages presented here mark the fundamental steps Wherever the mind encounters problems too difficult to solve with the functional instruments it has at its disposal, the process of comprehension recurs to the same mechanisms long left behind on other levels

A continuous process of interaction may be seen, new adaptations become possible with progressive mental development, and mental development progresses with new experiences The evolution of notions is one of the aspects of this development By relating the results of the present study to those of previous studies, it has been possible to show that the development of the notion of space is characteristic of the development of all notions, and in fact, of all instruments of thinking. Notions are not forms constructed once during the history of the mind and then applicable to all situations They are on the contrary the correlates of activity at various levels They are the instruments which the mind renews and recreates during its activity Their creation follows similar laws at all levels Their beginnings have to be sought in relation to the first development of sensory-motor activity This study has shown how they are newly constructed on a level of *practical* activity in a period between the first *sensory-motor* development and the coming of pure *reflective* thought They have been seen reappearing in a new constructive process at the level of reflective thought itself.

Besides this *theoretical* aspect, it is appropriate to mention possible *practical* applications for a study of this kind.

It has been shown in the last few years that in general the genetic psychology of reasoning contributes valuable instruments to the practice of psychodiagnosis By studying the functioning of the thought process rather than only its results, many diagnosticians have found themselves enabled to detect deviations and abnormalities by comparison with the normal genetic sequences. A genetic study of reasoning can provide valuable data in terms of general intellectual level. The study of complicated processes has proved especially useful from the standpoint of psychodiagnosis because of the

fact that complicated processes are the more sensitive ones. As might be expected, processes such as those involved in the evolution of notions like space and time are more readily impaired than simple ones.

Deviations from normal mental activity may be suspected when solutions in tests of the kind employed in the present investigation differ from those normally found. Although the data presented here are not sufficient to establish age norms the results, even in their preliminary form, may prove useful in a qualitative mental diagnosis.

Tests of the kind employed in the present investigation present still another useful feature. Their concrete set-up allows behavior studies of various kinds to be made. They yield information concerning steadiness of conduct, impulsiveness, emotionality, reactions to success and failure, etc. This report may make valuable contributions to personality diagnosis. The genetic study of normal children provides us with tools to distinguish between traits which are to be regarded as within the normal range of individual differences and those exceeding those limits.

Following the theoretical as well as the practical aspects described here, further studies might well point in two directions.

In pursuance of theoretical interests investigations should in the first place be deepened, and should include a more subtle analysis of the evolution of the notion of space, especially in studying problems which are on the borderline between the notions of space and time (for example problems concerned with seriation, movement, etc.)

In pursuance of practical interests investigations should be conducted with a greater number of subjects and with more standardized experimental situations than was possible in this preliminary approach.

G. Summary

Sixty-three children between the ages of 18 months and 5½ years were examined. Several experimental situations were devised. Comprehension of the spatial relations between objects were studied: (*a*) the fitting together of forms, (*b*) the comprehension of a moving object in relation to other objects. Comprehension of the individual's own shifts of position were studied, (*a*) their coordination with a freely movable object, (*b*) their coordination with an object whose movements are regulated.

Three genetic stages were distinguished, showing a development from a practical to a subjective and an objective notion of space. A comparable genetic sequence occurs on the reflective level.

These genetic trends are useful in psychodiagnosis for discussing and appraising individual deviations

References

1 Alpert, A The solving of problem situations by preschool children *Teach Coll Contrib Educ*, 1928, No 323

2. Buhler, C, & Hetzer, H Kleinkindertests Leipzig Barth, 1932

3. Gesell, A The Mental Growth of the Preschool Child New York Macmillan, 1925.

4 Gesell, A, & Thompson, H Infant Behavior Its Genesis and Growth New York McGraw-Hill, 1934

5 Guillaume, P, & Meyerson, I Recherches sur l'Usage de l'Instrument chez les Singes I. Problème du Détour *J de Psychol*, 1930, 27, 177

6 ——— Recherches sur l'Usage de l'Instrument chez les Singes VI Intermédiaire lié à l'Object *J de Psychol*, 1931, 28, 481

7 Kohler, W Intelligenzpruefungen an Anthropoiden Berlin Springer, 1921

8 Mathesons, E. A study of problem solving behavior in preschool children *Child Devel*, 1931, 2, 242-262

9 Meyer, Edith Ordnen und Ordnung bei drei-bis sechsjaehrigen Kindern *N Psychol Stud*, 1934, 10, 3.

10 ——— La Representation des Relations Spatiales chez l'Enfant Cahiers de Pédagogie et de Psychologie de l'Enfant Genève Univ Genève, 1935

11 Piaget, J La Naissance de l'Intelligence chez l'Enfant. Neuchatel-Paris Delachaux et Niestlé, 1936

12. ——— La Construction du Réel chez l'Enfant Neuchâtel-Paris Delachaux et Niestlé, 1937

13. Poincaré, H La Valeur de la Science Paris. Flammarion

14. Rey, A L'Intelligence Pratique chez l'Enfant Paris Alcan, 1935

15 Rey, A, & Lambercier, M Contribution a l'Etude de l'Intelligence Pratique chez l'Enfant *Arch de Psychol*, 1935, 26, 1-59

Clinic of Child Development
14 Davenport Avenue
New Haven, Connecticut

The Journal of Genetic Psychology, 1940, 57, 153-163

STUDIES IN HUNGER II THE RELATION OF GASTRIC DENERVATION AND DIETARY SUGAR TO THE EFFECT OF INSULIN UPON FOOD-INTAKE IN THE RAT*[1]

Department of Psychology, Harvard University

C T MORGAN AND J D MORGAN

A number of factors may determine the amount and kind of food ingested by man or animals Some of these conditions may be enumerated as follows (see 15, 22) contractions of gastric musculature (2, 3, 4), food-deprivation without hunger pangs (4, 9), dietetic deficiency accompanied by demands for special foods (7, 8), 12, 21), previous conditioning to specific foods (5), the excitation of sensory receptors associated with the ingestion of food (2), the amount of food presented to the organism (1), social stimulation frequently involved in eating (1), and finally, the health of the organism (4).

Attempts to classify the determiners of food-seeking behavior into simple categories have yielded the concepts of "hunger" and "appetite" According to Cannon (2, p 20), hunger is a "primitive, elemental sensation," experienced as a "dull ache or gnawing pain" which originates in stomach contractions, and appetite "arises from the experience of previous pleasures" and refers to "wishing, longing or yearning for something specially desirable" Carlson (4) agrees in essence with Cannon's idea of hunger but holds that appetite is the expression of an inherited desire for food rather than the composite of learned reactions and anticipatory images associated with eating (cf, 2, pp. 20-21).

Although the concepts of Cannon and of Carlson seemed to be supported by the introspective observations of man, when they are applied to animals, certain difficulties are encountered For one thing, in certain animals (for example, insects, salmon, birds, herbivorous mammals), feeding behavior is obviously not determined

*Received in the Editorial Office on July 29, 1939

[1]This study was carried out in the Psychological Laboratories of the University of Rochester

by gastric motility and the emptying of the alimentary tract (see 4). Furthermore, gastrectomized rats (19) and human individuals (see 9) may possess what appears to be a normal desire for food. Finally, the presence of general restlessness and an increased level of activity seems just as typical of the demand for food in animals (4) and man (20) as is motility of the stomach.

A definition of hunger which takes account of the above objections to a strictly local theory of hunger, and which allows for the as yet limited knowledge of its nature, seems to be needed. The authors offer the following descriptive definition: hunger consists of the physiological conditions within an organism, not determined in any significant way by previous learning or concomitant exteroceptive stimuli, which give rise to food-seeking and food-taking responses. The best measure of hunger, so considered, would seem to be the amount and rate of food-ingestion under conditions which hold constant the factors related to past learning and exteroceptive stimulation. It is in this frame of reference that the problems studied in this and a previous investigation (13) have been formulated to the end of unearthing facts of significance concerning the physiological basis of hunger.

Problem

Many researches have shown that the injection of insulin produces marked increases in gastric motility (10, 14, 16, 17, 18). It has been shown also that preceding such insulin hypermotility there occurs a phase of gastric hypomotility (14). The effect of insulin on the food-intake of rats therefore seemed a question which might be profitably investigated in relation to the general question of the physiological basis of hunger, and in relation also to the specific problem of the part that gastric motility may play in the mechanism of hunger.

A previous study by the authors (13) yielded the following results concerning the effect of insulin-injection upon the food-intake of rats. The injection of insulin can produce either an increase or decrease in the rate of eating. Which of the two possibilities is realized depends upon the interval which elapses between the injection and the time at which feeding is begun. An increase in food-intake accompanies the use of an interval of 15 minutes and a decrease fol-

lows the use of 30- or 60-minute intervals The variable effect of insulin upon food-intake therefore corresponds in some degree with known relations between insulin-injection and gastric motility

With reference to the general problem of the physiological basis of hunger, the question which arose from the above results was this Does the change in hunger-strength brought about by insulin have its basis in changes in gastric motility or in some other physiologic modification, for example hypoglycemia? It seemed that this question might be answered experimentally by comparing the effects of insulin on normal and vagotomized animals[2] The present paper is a report of the results secured in the investigation of this particular problem

Method

Fourteen albino rats, approximately five months of age, were trained to depress a lever to obtain food in a modified Skinner-apparatus (6) These rats were then split into two groups Each animal of one group was matched with a litter-mate of the same weight in the other group

One group of seven rats was subjected to the following operation performed under deep ether anesthesia An incision, about one inch in length, was made in the rat's abdomen just below the xiphisternum The intestines and stomach were drawn out through the incision and the major arteries which enter the stomach were tied The main vagal trunks, which could be clearly seen, were then sectioned somewhat below the diaphragm. In order to make certain that small branches of the vagus nerve did not remain intact, the stomach was freed entirely from its supporting mesenteries and vascular supply Following this the stomach and intestines were replaced in the abdomen and the incision closed with surgical silk

Recovery from the operation was relatively rapid and no forced feeding was required. Indeed, some of the rats were observed to resume feeding within a few hours after the operation

All 14 rats were retrained in the feeding apparatus two or three days after the operations on the one group had been completed From

[2]It has been shown in previous researches (see 16) that the increase in gastric motility occasioned by insulin-injection is eliminated by resection of the vagus nerve

this time on, all animals secured all of their food in the apparatus None of the records were considered a part of the experimental data however until the retraining and feeding schedule had been maintained for at least four days Daily sessions were exactly one hour in length and, for a given rat, were begun at the same time every day A record of the hour's feeding curve was secured on waxed paper[3] attached to a slowly moving kymograph-drum Every depression of the lever was rewarded by a "shot" of food approximately one-tenth of a gram in weight Except in a special case explained later in the paper, the diet used throughout the entire study was a granular mixture of the following proportions by weight: yellow corn meal, 50 per cent; Purina dog chow, 30 per cent; and sucrose, 20 per cent

As in the previous study (13), an odd-even method of self-control was used On one day the rats were injected with insulin and on the next day an equal amount of sterile saline solution was injected. Thus the saline days provided control records with which the records of the insulin days could be compared The dosage of insulin for each rat was 01 cc of 40U Iletin insulin per 10 grams weight

Both normal and vagotomized animals were subjected to three experimental series which differed from each other only in the amount of time that was allowed to elapse between the time of injection of the insulin or saline and the time that the rats were placed in the box Series I employed an interval of 0-minutes, that is, the rats were placed in the feeding box immediately after they were injected; in Series II, the interval was 15 minutes, in Series III, it was 30 minutes The different experimental series were carried out in the order named

In each experimental series, about 30 pairs of feeding curves were secured For each curve following insulin injection, a control (saline) curve was obtained. When either a normal or an insulin curve was lost due to technical difficulties, as was sometimes the case, the corresponding control or insulin curve for the day before or after the curve in question was discarded.

The curves were translated into numerical scores by means of a

[3]Graciously provided by the Waxon-Carboff Corporation of Rochester, New York

calibrated scale. The number of responses for each 10 minutes of the hour's feeding was read from the scale and recorded. The obtained scores were converted into percentages of the mean number of eating responses on the control days of each series. From these percentages, the mean rate of response for the six 10-minute intervals and the standard errors of these means were determined. The correlation between an insulin score and the corresponding control scores of the following day was computed according to Pearson's method of moments. From these values, the statistical reliability of the differences of the means of control and insulin days was computed for both normal and vagotomized rats in terms of the ratio of the obtained difference to its standard error.

RESULTS

Table 1 shows the results of the three experimental series conducted with normal animals and Table 2 gives comparable data for

TABLE 1

RELIABILITY OF DIFFERENCES OF CONTROL AND INSULIN MEAN PER CENT RESPONSES PER TEN MINUTE INTERVAL FOR NORMAL RATS

Minutes	0-10	10-20	20-30	30-40	40-50	50-60	Total
0 Minutes ($N = 33$)							
Saline	26.21	19.39	17.24	14.97	12.24	10.15	101.06
Insulin	26.88*	18.94	15.61	14.18	10.24	7.64	92.42
Diff	.67	.45	1.63	.79	2.00	2.51	8.64
σ_{diff}	1.39	.73	.86	1.02	1.01	1.07	4.16
D/σ_{diff}	.48	.62	1.91	.78	1.97	2.36	2.08
15 Minutes ($N = 33$)							
Saline	27.58	18.91	15.64	13.15	12.91	11.36	100.66
Insulin	25.39	18.09	15.45	13.94*	12.91	11.30	97.51
Diff	2.19	.82	.19	.79	.00	.06	3.15
σ_{diff}	.92	.64	.81	.93	.75	1.08	3.10
D/σ_{diff}	2.36	1.27	.23	.81	.00	.05	1.02
30 Minutes ($N = 28$)							
Saline	24.86	18.75	16.57	15.24	13.25	11.18	100.40
Insulin	19.71	16.18	15.29	13.11	12.21	11.11	87.50
Diff	5.15	2.57	1.28	2.13	1.04	.07	12.90
σ_{diff}	1.02	.83	.61	.52	.97	1.07	2.76
D/σ_{diff}	5.07	3.10	2.12	4.13	1.07	.07	4.67

*Means which are higher than corresponding control values

TABLE 2

RELIABILITY OF DIFFERENCES OF CONTROL AND INSULIN MEAN PER CENT RESPONSES PER TEN MINUTE INTERVAL FOR VAGOTOMIZED RATS

Minutes	0-10	10-20	20-30	30-40	40-50	50-60	Total
0 Minutes ($N = 31$)							
Saline	33.87	22.26	19.77	12.97	6.71	4.19	100.32
Insulin	38.71*	29.32*	21.48*	14.77*	9.52*	4.65*	118.79*
Diff.	4.84	7.06	1.71	1.80	2.81	.46	18.47
σ_{diff}	2.07	1.83	1.73	.99	.45	1.11	5.55
D/σ_{diff}	2.34	3.86	.99	1.81	6.22	.41	3.33
15 Minutes ($N = 31$)							
Saline	27.65	23.71	20.71	14.61	9.19	4.10	101.12
Insulin	28.39*	22.42	18.97	15.26*	13.26*	9.35*	108.69*
Diff.	.74	1.29	1.74	.65	4.07	5.25	7.57
σ_{diff}	1.29	.80	.60	.78	1.41	1.67	4.01
D/σ_{diff}	.58	1.61	2.93	.83	2.88	3.16	1.89
30 Minutes ($N = 29$)							
Saline	32.97	24.59	19.72	12.97	6.93	3.72	100.93
Insulin	27.78	24.35	20.72*	16.62*	13.79*	10.14*	113.79*
Diff.	5.19	.24	1.00	3.65	6.86	6.42	12.86
σ_{diff}	1.87	.93	.90	1.27	1.42	1.11	4.51
D/σ_{diff}	2.78	.74	.11	2.87	4.82	5.77	2.85

*Means which are higher than corresponding control values.

the vagotomized animals. The differences in rate and also the reliability of these differences between control and insulin feeding curves are given for each 10 minutes of the hour-period and for the total amount of food-intake during the hour.[4] The tables reveal striking differences in the effect of insulin upon food-intake in normal and vagotomized rats. Almost without exception, the feeding rate of normal rats is depressed following insulin and, on the other hand, the feeding rate of vagotomized rats is in practically every case accelerated.

A common characteristic of the effect of insulin upon both normal

[4]It is to be noted that the control mean totals in Tables 1 and 2 are not exactly 100, the value which should obtain in view of the use of per cent-scores, and that the means for the ten minute-intervals when totaled do not give the same values as the mean total directly computed. These discrepancies are attributed to the use of grouped data for the calculations.

and operated rats is the cyclic nature of the changes in feeding rate (see Tables 1 and 2) These cycles of facilitation and depression (or lack of facilitation) are plainly not artifacts or chance variations, for they appear consistently in the experimental series of this and our previous study (13) Figure 1 in which normal rats on

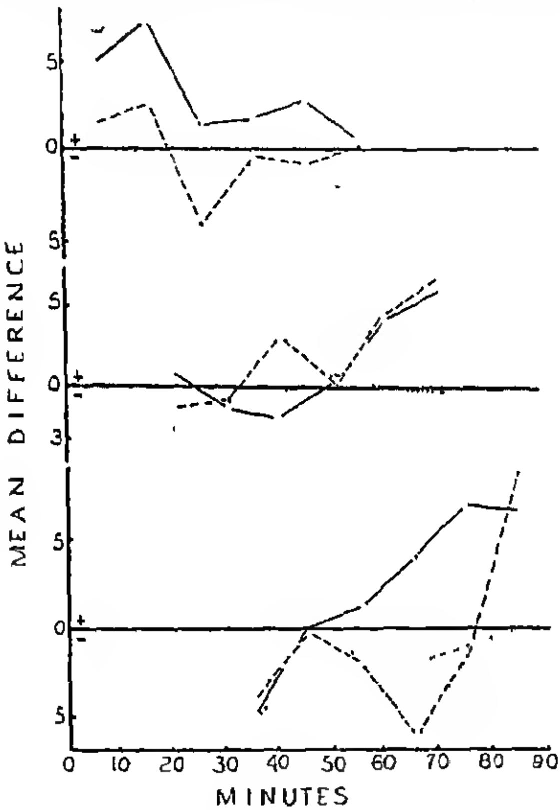

FIGURE 1

THE MEAN DIFFERENCE OF INSULIN AND CONTROL EATING RATES FOR NORMAL RATS ON SUGAR FREE DIET (DASHED LINES), NORMAL RATS ON HIGH SUGAR DIET (DOTTED LINES), AND VAGOTOMIZED RATS ON HIGH SUGAR DIET (SOLID LINES)

Values are plotted for each 10 minutes of an hour's feeding period. The upper graph is obtained in Series I, the middle graph in Series II, and the lower graph in Series III For further explanation, see text.

low sugar diet,[5] normal rats on high sugar diet, and vagotomized rats on high sugar diet are compared with respect to the effect of insulin upon eating rate, demonstrates this point

From the data presented in Tables 1 and 2 and in Figure 1, the following observations may be made. (*a*) Food-intake is increased immediately following the injection of insulin (zero to 20 minutes after injection) This effect is quite large in vagotomized rats, less pronounced in normal rats on a sugar-free diet, and practically negligible in rats with high sugar diet. (*b*) Following the initial excitation there is a drop from the insulin supernormal rate (20 to 50 minutes after injection) This depression is greatest for the rats on the sugar-free diet, somewhat less for the animals on high sugar diet, and still less for the vagotomized rats (*c*) During the latter part of the period after injection (50 to 90 minutes) there is a large and significant increase in the number of feeding responses in vagotomized rats and normal rats on the sugar-free diet On the other hand, the normal rats eating the high sugar diet recover from the preceding depression but show no such supernormal rate of eating. (*d*) In general, insulin increases the amount of food eaten by vagotomized rats, it has a variable effect upon rats kept on a low sugar diet, and usually depresses normal rats maintained upon a high sugar diet

The substitution of sugar for the gelatin used in the diet of the first study (13) seems to have produced significant effects As a test of these conclusions concerning the relation of sugar-content of the diet to the effect of insulin upon food-intake, a check experiment was carried out The diet of the group of normal rats used in the experiments described above was modified to the extent of substituting yellow corn meal for the sugar The diet was in this way made comparable to that used in our first study (13) with respect to sugar content Series II was then run again. Table 3 shows the results of this check and compares them with the results previously reported Group I refers to the rats used in the earlier study and Group II consists of the rats studied in the present research As Table 3 shows, the repetition of the experiment yields almost iden-

[5]These data are those reported in greater detail in our first study (13) The diet of that study differed from the diet of the present study in that gelatin was used instead of sugar

TABLE 3

A Comparison of the Effect of Insulin Upon the Feeding Curves of Two Groups of Normal Rats on the Same Diet (Sugar-Free)*

Minutes		10	20	30	40	50	60
Group I							
($N = 187$)	Saline	23.08	43.71	62.68	79.10	91.13	99.79
($N = 44$)	Insulin	22.04	42.00	63.68	80.41	96.77	111.55
	Diff	1.04	1.71	1.00	1.31	5.64	11.76
	σ_{diff}	1.29	2.21	2.93	3.10	2.65	3.80
	D/σ_{diff}	.80	.77	.34	.42	2.13	3.09
Group II							
($N = 34$)	Saline	24.71	43.56	62.68	76.62	89.65	100.77
($N = 34$)	Insulin	25.47	44.30	62.97	79.06	94.12	111.76
	Diff	.76	.74	.29	2.44	4.47	10.99
	σ_{diff}	.83	1.60	2.04	2.31	2.72	3.38
	D/σ_{diff}	.92	.46	.14	1.06	1.65	3.25

*The interval between injection and feeding is 15 minutes.

tical results. This experimental check accordingly confirms the conclusions drawn above concerning the importance of the sugar-content of the diet in the effect of insulin upon hunger, and lends assurance of the reliability of the procedures and of the results of our present methods.

Discussion

The conclusion that the feeding rate of insulin-injected rats varies from the normal in a cyclic manner seems to be clearly supported by the data. The authors do not wish to speculate at this time as to the physiological basis of such cyclic effects. Profound endocrine and metabolic disturbance can be expected to result from the substantial doses of insulin that were used in this study, and opposing physiological processes which possess definite periodicity can well be expected.

Of more specific value in the quest for immediate determinants of the drive for food is the fact that vagotomized animals are made more "hungry" by insulin than are normal animals. This experimental fact is probably related in a fundamental way to the second important result of the present study, that the amount of sugar in the diet determines to some extent the effect of insulin upon hunger. Vagotomized animals exhibit a reduced capacity to convert and ab-

sorb disaccharides (see 4) so that the characteristic effect of insulin upon such animals may be due directly to low sugar storage and utilization occasioned by the vagotomy. The authors are at present conducting further work on this particular problem.

Summary and Conclusions

Since insulin-injection was found, in a previous research, to produce significant changes in hunger, as measured by the rate and number of food-taking responses, insulin-injection has been used as an experimental tool in the study of the general problem of the nature of hunger. The food-intake of insulin-injected rats was measured in a Skinner-type apparatus relative to the effect of diminishing gastric motility by resection of the vagus nerve and in relation also to the effect of sugar in the diet. It was found that denervation of the stomach not only failed to eliminate the augmentation of hunger due to insulin-injection but that such an operation actually enhanced the effect. It was found also that reduction in the amount of sugar in the diet served to increase the effect of insulin in strengthening the hunger drive. The results are interpretated to indicate that nutritive conditions of the organism may be of more direct significance in the control of hunger than are stomach contractions.

References

1. Bayer, E. Beitrage zur Zweikomponententheorie des Hungers (Versuche mit Huhnern). *Zsch. f. Psychol.*, 1929, **112**, 1-54.
2. Cannon, W. B. Digestion and Health. New York: Norton, 1936.
3. Cannon, W. B., & Washburn, A. L. An explanation of hunger. *Amer. J. Physiol.*, 1912, **29**, 441-454.
4. Carlson, A. J. The Control of Hunger in Health and Disease. Chicago: Univ. Chicago Press, 1916.
5. Elliott, M. H. The effect of change of reward on the maze performance of rats. *Univ. Calif. Publ. Psychol.*, 1928, **4**, 19-30.
6. Hanford, H. M., & Morgan, C. T. A simplified lever-operated device for conditioning studies in rats. *J. Gen. Psychol.*, 1939 (to appear).
7. Harlow, H. F. Food preferences of the albino rat. *J. Genet. Psychol.*, 1932, **41**, 430-438.
8. Hausman, M. F. The behavior of albino rats in choosing foods and stimulants. *J. Comp. Psychol.*, 1932, **13**, 279-307.
9. Hoelzel, F. Central factors in hunger. *Amer. J. Physiol.*, 1927, **82**, 665-671.

10 LA BARRE, J, & DESTREE, P L'influence des variations glycemiques sur la motilite gastrique *C R Soc Biol Paris*, 1930, 103, 532-534
11. M'CREA, E D, M'SWINEY, B H., & STOPFORD, J B Effects of the section of the vagi nerves on the motor activity of the stomach *Quart J Exper Physiol*, 1926, 16, 195-203
12 MASLOW, A H Appetites and hungers in animal motivation *J Comp Psychol*, 1935, 20, 75-83
13 MORGAN, C T, & MORGAN, J D Studies in hunger I The effect of insulin upon the rat's rate of eating *J Genet Psychol* (to appear).
14 MULINOS, M G The gastric hunger mechanism IV. The influence of experimental alterations in blood sugar concentration on the gastric hunger contractions *Amer J Physiol*, 1933, 104, 371-378
15. MURSELL, J L Contributions to the psychology of nutrition I Hunger and appetite *Psychol Rev*, 1925, 32, 317-333.
16. QUIGLEY, J P, & TEMPLETON, R D Action of insulin on the motility of the gastro-intestinal tract IV Action on the stomach following double vagotomy *Amer J Physiol*, 1930, 91, 482-487
17 REGAN, J F The action of insulin on the motility of the empty stomach *Amer J Physiol*, 1933, 104, 90-94
18 TEMPLETON, R D, & QUIGLEY, J. P The action of insulin on motility of the gastro-intestinal tract. *Amer J Physiol*, 1930, 91, 467-474
19 TSANG, Y C Hunger motivation in gastrectomized rats *J Comp Psychol*, 1938, 26, 1-17
20 WADA, T. An experimental study of hunger in its relation to activity *Arch of Psychol*, 1922, 8, No 57 Pp 65
21 WILDER, C E. Selection of rachitic and antirachitic diets in the rat *J Comp Psychol*, 1937, 24, 547-577
22 YOUNG, P T Preferences and demands of the white rat for food. *J. Comp Psychol*, 1938, 26, 545-590

Department of Psychology
Harvard University
Cambridge, Massachusetts

The Journal of Genetic Psychology, 1940, 57, 165-191

PERSEVERATION OF NON-REWARDED BEHAVIOR IN RELATION TO FOOD-DEPRIVATION AND WORK-REQUIREMENT*

Department of Psychology, University of Tennessee

PAUL M. FITTS[1]

A. INTRODUCTION

The rate of occurrence of an objectively defined behavior sequence can be shown to vary under the influence of numerous experimental conditions. Although results obtained from a study of an isolated segment of behavior cannot be generalized to cover all forms of activity, nevertheless intensive study of simple behavior sequences is desirable because it makes possible an analysis of the effects of several simultaneously acting variables and an evaluation of their inter-related influences.

In the investigation here reported a study was made of the influence of two experimental variables upon the frequency of occurrence of a simple bar-pressing response. Forty white rats were given an equal number of reinforcements in making this response. Food-reward was then removed and at intervals over a period of seven months determinations were made of the rate at which the animals continued to press the bar. Four experimental conditions, each consisting of a different combination of food-deprivation and work-requirement, were used in securing data on the interrelated influence of these two factors upon the persistence of response. The findings of the present study supplement the facts which are already known about the influence of single variables upon the rate of response. The period of time during which animals were studied in the present investigation, moreover, was greater than that covered by any similar study.

Among the recent studies which bear on the present problem are those of Bousfield (1), Bousfield and Elliott (2), Skinner (17),

*Accepted for publication by Leonard Carmichael of the Editorial Board, and received in the Editorial Office on August 5, 1939.

[1]The writer is indebted to Dr. Leonard Carmichael, Tufts College, for counsel in planning and conducting this investigation. The study was carried out in the Laboratory of Psychology of the University of Rochester.

Young (26) and Fitts (10) These investigators have all secured quantitative results which show that the frequency of rewarded food-seeking response varies in a predictable manner with changes in the period of deprivation and also with the amount of food ingested Skinner (19), Youtz (27), Williams (25), and Fitts (10) have studied the frequency of non-rewarded response or "resistance to extinction" in relation to deprivation or in relation to the number of training reinforcements Their results are in general agreement The rate at which non-rewarded responses are made during short periods of time is a function of the number of responses which were previously rewarded and of the condition of deprivation existing in the animal at the time determinations are made

The rate at which animals make responses is influenced also by the nature and difficulty of that particular behavior Intense stimulation, such as shock, and excessive work, such as might be encountered in traversing long pathways or crossing other obstructions, have been found to reduce the frequency of response.

It is more difficult to determine the manner in which various combinations of such factors as training, motivation, and difficulty of response collectively modify behavior than it is to determine the variations produced by any one of these factors taken singly Because of this difficulty the nature of such interrelated influences has been postulated from indirect evidence as often as it has been studied through direct experimentation. One such general inference is that several mechanisms are needed to account for the complexity of the relationships between motivation and behavior Warden, in describing the activity of rats after long starvation periods, stated that, "One is tempted to say that the hunger drive is as strong as ever and that the reduction in crossings is due to decreased capacity to resist the shock" (23, p. 70). It is assumed, therefore, that as deprivation increases, the drive for food tends increasingly to facilitate behavior while at the same time the lessened resistance to shock tends increasingly toward inhibition of activity Interpretations of a similar nature have been suggested for many situations where activity or rate of response at first increases with deprivation, then reaches a peak and begins to decline. In several studies of the rate of eating (2, 10, 14), and of the level of general bodily activity (13, 16, 22), such phenomena have been reported The decrease in rate of performance, it must be noted, takes place in spite of the obviously increasing need for nourishment

Stone (21) has made a direct investigation of the influence of work upon frequency of response. He studied the perseveration of rats in digging through sand and found a marked continuation of activity after the removal of food-reward. This result would indicate that a certain minimum amount of work may be done regardless of the presence or absence of food-reward. At times animals have been found to respond less often when the need for food was increased, while at other times they have been found to persist in working when no food was secured.

A consideration of such findings makes it obvious that the rate at which an animal performs a given behavior sequence is determined by the combined influence of many factors. The contributing influence of each of these variables when acting in combination with others should be evaluated in any fundamental study of the relation between motivation and behavior. Especially is a knowledge of their effect upon learning and unlearning important. In the present investigation this problem was approached through a study of the influence of two variables upon the persistence of animals in making a simple behavior sequence. Some supplementary data were also secured on the extent to which an alternative or competing activity would affect the response-decrement.

B. Apparatus

The apparatus and experimental techniques used in this investigation are similar to those developed by Skinner (20). Animals were trained to press a small lever by rewarding them after each response with a small portion of food. Responses were recorded automatically. The animals were not disturbed by the experimenter at any time during the periods of experimentation.

Figure 1 shows a general view of the animal compartment and lever. The sides of the compartment are made of metal while top and bottom are of wire mesh. A bar, six millimeters in diameter and six centimeters from end to end, extends through a panel at one end. This bar is mounted 10 centimeters above the floor and extends two centimeters into the compartment. It can be moved downward for a distance of two centimeters and is pivoted and counterweighted on the outer side of the panel in such a way that the force required to depress it can be varied to suit the needs of the experiment. On the outer end of the lever is a Mercoid mer-

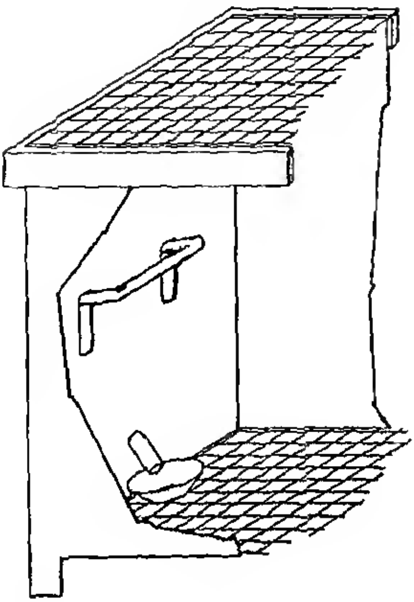

FIGURE 1
ANIMAL COMPARTMENT, SHOWING LEVER AND FOOD-PAN

cury switch, which is adjusted so that an electrical circuit is completed whenever the bar is moved downward the full distance of two centimeters and contact broken after the bar has returned most of the distance toward its original resting position. Immediately below the lever is a small food-pan, to which a glass tube leads from the food-releasing device above.

The food-releasing mechanism is shown in Figure 2. The solenoids (*M*) which furnish motive power for this food-release are placed in circuit with a 6-volt storage battery and the mercury switch. The device is thus activated each time the animal pushes the lever downward. The bin (*C*) is filled with powdered food. In the bottom of this bin is a cylinder (*A*) drilled with eight small holes (*B*) which measure 6 mm. by 6 mm. The lower half of the cylinder is fitted to a bushing which is so designed that as the cylinder is rotated by the pull of the solenoids the food, which falls into the upper holes or food-pockets, is carried around and emp-

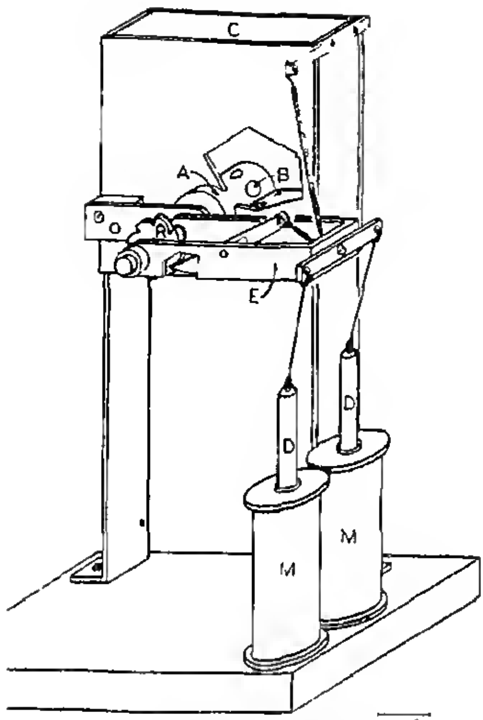

FIGURE 2
Food Releasing Device

tied, one pocket at a time, into the upper end of the glass tube leading to the animal compartment

The entire apparatus, including animal compartment and food-releasing mechanism, was assembled within a light-proof box with four-inch sound-shielded walls This larger box was ventilated by a compressed air system The recording drums were located at some distance from the rest of the apparatus. Animals were never observed directly or otherwise disturbed at any stage of training or subsequent testing.

Records of bar-pressing activity were secured by means of a pen and ink recording system A series of kymograph drums, mounted on a single shaft, were turned by a Telechron motor at the rate of

one revolution in six hours. Above the drums were mounted recording pens adjusted to move at right angles to the direction of rotation of the drums. An electrically activated wheel and rachet device, in circuit with the mercury switch and food-release, allowed the pen to move a distance of 1 mm across the paper covered drum each time an animal made a response The records thus secured were in the form of summated time-response curves (see Figure 3)

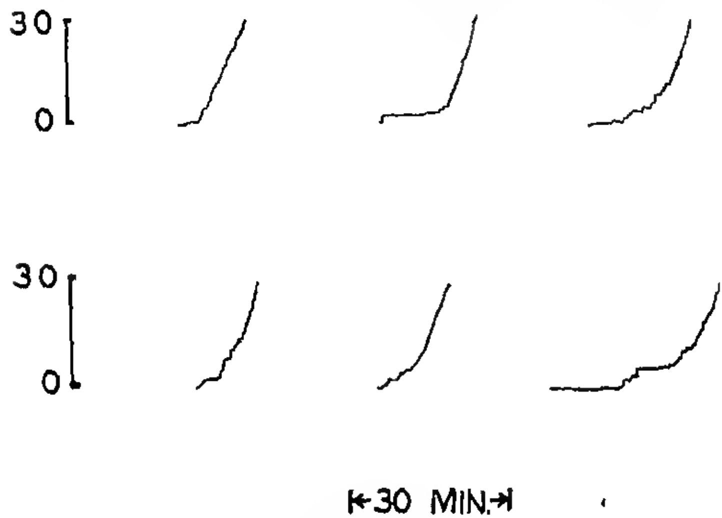

FIGURE 3
SIX CHARACTERISTIC LEARNING CURVES

An impulse counter was used as a check to determine the exact number of responses made during the complete experimental period Throughout all of the investigations two duplicate experimental set-ups were employed and two animals were experimented upon at the same time This procedure made it possible to control various external influences by having animals from two different groups working at the same time

During the course of the experiment a rotating wheel was added to each experimental apparatus These running wheels were 30 cm in diameter and 10 cm. in width. One side was connected through a small door with the end of the animal compartment opposite the

lever The sides of the running wheels were of metal, the floors of one-half inch wire mesh The wheels were mounted on bearings and revolved quite easily Any movement of the wheel was recorded by means of an eccentric and a light thread which led to a marking pen on the recording drum.

C Procedures

Forty male albino rats of an inbred laboratory stock were employed in the investigation These animals were secured from 12 litters They were weaned at three weeks of age and placed in small living cages. Training began when they were approximately 50 days of age

Each animal was placed on a feeding schedule two weeks before the beginning of training. Food (Purina dog chow) was placed in racks in the living cages at the same time each day and the uneaten food removed two hours later On this feeding schedule the animals continued to grow and to gain weight

The first step in training was to teach the animals to go to the food-pan at the sound of the food-releasing mechanism, so that they would secure the reward immediately after making the first bar-pressing response In order to do this each animal was allowed to eat in the animal compartment during one feeding period, with the lever detached At a corresponding hour on the succeeding day the animal was placed again in the compartment and 0 05 gram of food was delivered automatically every 30 seconds by a timing device temporarily inserted in the circuit After 30 minutes, or after receiving 60 automatic food-deliveries, the animal was returned to its living cage and given additional food The same procedure was repeated on the following day This preliminary training was similar to that employed by Youtz (27) and proved successful in that it resulted in very rapid learning when the animals were allowed to initiate their own food delivery

The next step in training was to teach the animals to press the bar in order to secure food The bars were mounted in the compartments and adjusted so as to require 10 grams of force to move them downward. The food-releasing device was connected. A trace of moist food was smeared over each bar (27) An animal was then placed in each compartment at the hour at which it was accustomed to eat, and allowed to remain until it had pressed the bar

exactly 30 times and secured 30 0.05 gram portions of food. After the 30th response each animal was removed from the experimental situation and returned to its living cage This completed the training At no time during the next seven months did any animal secure additional food-rewards in the experimental situation All 40 animals were trained in the course of a single week

The rate of learning during the training period was found to be very rapid The animals required between 5 and 30 minutes to reach the criterion of 30 responses counting from the first to the 30th response Figure 3 shows six typical learning curves These records show about the maximum variation in rate of learning

After the completion of training the 40 animals were divided into four smaller groups of 10 each All available data, including genetic background, age, weight, and speed of learning were considered in dividing the animals in such a way as to equate the four groups as nearly as possible

Determinations of non-rewarded behavior were begun on the seventh day after each animal had completed its training The following conditions of deprivation and work prevailed for each of the four groups during this first and all subsequent determination. *Group A-1,* 48-hours deprivation and 10 gm-cm of work, *Group A-2,* 48-hours deprivation and 100 gm-cm of work, *Group B-1,* 12-hours deprivation and 10 gm-cm. of work; *Group B-2,* 12-hours deprivation and 100 gm-cm of work

The amount of work was controlled by adjusting the counterbalance of the lever. The condition of deprivation was controlled by omitting the feeding period on the day preceding a determination in the case of animals in the "*A*" groups, and by giving an additional two-hour feeding period 12 hours before the time for experimentation in the case of "*B*" animals. All determinations on a given animal were made at the hour at which it had been accustomed to eating Two animals from different groups were tested at the same time and the sequence of determinations from hour to hour was such as to rule out any influence of diurnal changes in activity level, at least as far as the group results are concerned

The procedure followed in making the determinations of non-rewarded behavior was as follows The food-release was disconnected, so that animals no longer could secure food nor hear the sound of the food-releasing mechanism Animals were placed suc-

cessively in the experimental compartments, where they were allowed to remain during the test period. Eleven determinations in all were made on each animal. The first six were made at intervals of one week, the seventh after four weeks, the eighth after six weeks, and the remaining three at intervals of four weeks. The first six of these experimental periods were one hour in length, the last five periods were each four hours in length. Periodic feeding was continued during the seven months covered by the investigation with the exception of four weeks between the 6th and 7th determinations. In this case the regular feeding schedule was resumed two weeks prior to the seventh experimental period. On experimental days the animals were fed one hour after the end of the test period.

If an animal made a response during the last few minutes of the experimental period it was allowed to remain a short time longer, until there was a period of quiescence, before it was removed. This was done in order that removal from the compartment would be less likely to serve as reinforcement to the last response. Only responses made during the pre-determined experimental period were tabulated, however.

At the end of the ninth determination, or after five months, the tendency toward perseverance in the bar-pressing behavior had become stable in the constant environment furnished by the animal compartment. The question then arose as to whether this perseveration would continue if the environment was modified by the introduction of an alternative activity. To answer this question a running wheel, as previously described, was mounted in each experimental box. In the following weeks half of the animals in each group were allowed access to these running wheels during the experimental period and half were not.

In addition to the data secured on the regular experimental animals, records were taken on a group of 15 control animals. This control group was composed of male rats of an age and weight and genetic background similar to the experimental animals. They were given all of the preliminary training, including the automatic feeding in the animal compartment, but they were not given training in pressing the bar. Tests were given one week later to determine the number of bar-pressing responses which would be made without training. The determinations on the control group were made under conditions which should have led to a maximum level of activity,

TABLE 1

Number of Non-Rewarded Responses Made by Individual Animals During Each Experimental Period

Animal	Period 1	2	3	4	5	6	7 1hr	7 2hr	7 3hr	7 4hr	8 1hr	8 2hr	8 3hr	8 4hr	9 1hr	9 2hr	9 3hr	9 4hr
Group *A*-1																		
34	155	116	76	114	122	118	201	110	87	59	190	58	219	105	15	13	0	0
40	150	57	75	53	49	44	43	35	15	15	20	26	30	12	31	31	26	8
42	132	90	59	68	35	48	25	26	18	19	48	55	57	18	19	13	22	5
10	104	39	22	34	41	45	37	8	3	12	43	71	54	34	25	33	22	28
38	95	55	76	112	107	70	20	48	17	12	33	72	19	90	76	57	57	34
26	90	36	43	20	42	29	26	52	45	24	8	87	57	10	99	42	35	18
14	86	98	49	145	78	44	78	114	38	33	57	95	41	21	30	28	19	11
30	75	22	13	19	42	82	31	32	34	21	60	42	27	34	23	10	63	23
22	61	38	26	38	40	18	22	81	32	9	75	135	57	35	24	19	2	11
18	51	30	33	22	14	12	6	17	33	20	65	59	51	23	29	41	25	18
Group *A*-2																		
27	118	41	44	33	34	49	43	44	29	22	80	22	42	17	24	53	17	38
35	80	94	40	88	70	41	53	42	66	15	24	17	47	4	17	6	5	2
39	78	16	1	26	31	7	0	16	35	17	12	25	15	21	13	7	7	23
11	69	47	20	31	35	43	37	41	44	20	20	5	20	27	26	30	16	6
19	59	32	12	44	11	26	2	1	0	1	1	0	8	6	12	3	2	1
31	53	13	6	18	10	49	16	7	7	5	26	12	0	20	88	56	0	41
23	51	92	28	31	53	39	41	32	19	15	9	34	10	57	11	2	1	6
43	43	44	25	15	60	23	31	38	21	23	11	9	17	22	25	1	8	0
15	35	4	3	8	22	16	45	47	46	23	60	14	74	19	23	42	12	19
47	19	20	63	41	26	9	3	29	18	13	10	9	9	4	4	4	0	2

TABLE 1 (*continued*)

Animal	Period 1	2	3	4	5	6	7 1hr	7 2hr	7 3hr	7 4hr	8 1hr	8 2hr	8 3hr	8 4hr	9 1hr	9 2hr	9 3hr	9 4hr
Group *B*-1																		
16	133	36	31	64	29	65	22	12	24	6	12	34	77	20	18	21	22	16
48	120	34	46	45	28	58	84	76	68	21	63	50	57	50	27	9	47	30
40	86	44	5	51	136	116	126	32	29	0	80	112	139	59	40	38	68	117
28	83	10	5	4	7	2	5	1	6	10	16	7	28	5	5	8	3	6
36	67	23	15	47	31	34	32	10	32	35	15	9	6	31	32	32	19	26
44	62	18	16	24	12	8	2	0	12	4	40	51	18	16	39	11	10	28
12	58	105	23	68	39	119	105	67	14	8	72	61	34	40	19	38	95	14
20	58	17	15	21	35	8	63	20	6	5	18	22	14	36	17	21	24	14
24	55	39	18	32	46	33	38	8	9	22	21	72	56	20	20	19	26	6
32	54	20	16	42	12	12	5	7	3	2	8	16	11	15	26	8	9	18
Group *B*-2																		
25	138	48	15	6	5	9	11	1	1	2	55	51	37	48	17	55	59	41
37	127	49	30	94	27	95	26	34	58	13	32	42	51	58	24	30	13	13
45	93	48	7	11	8	27	16	30	17	13	18	7	20	27	22	8	27	8
17	77	22	8	27	28	74	73	41	14	27	5	17	11	24	13	19	27	26
29	68	12	9	18	6	24	11	19	4	3	4	3	6	3	1	8	1	2
20	45	16	1	28	50	18	56	20	32	0	38	54	13	6	20	2	15	8
41	29	6	13	21	0	2	25	0	15	0	3	6	0	27	6	5	1	1
49	27	0	0	1	2	0	0	0	1	1	1	2	2	3	1	2	2	2
33	20	1	0	1	1	14	0	4	0	3	47	15	23	4	22	4	4	30
13	16	7	20	55	84	6	6	1	5	2	142	51	15	40	*	*	*	*

i.e., when they were doing only 10 gm-cm. of work at 48-hours of deprivation.

D. General Results on All Groups

The data on individual animals for the first 20 weeks of the investigation are presented in Table 1. These data show that the great majority of animals continued to make bar-pressing responses during this time even though food-reward was not secured. The smallest mean number of responses occurring during the first hour of any period was 25. This low level of behavior occurred during the third week and again during the 20th week. The mean number of responses for the 15 control animals, on the other hand, was 8±4.2 during the first test period. These control animals, as previously noted, were tested under conditions which should have led to a high level of random activity, but they had never received food

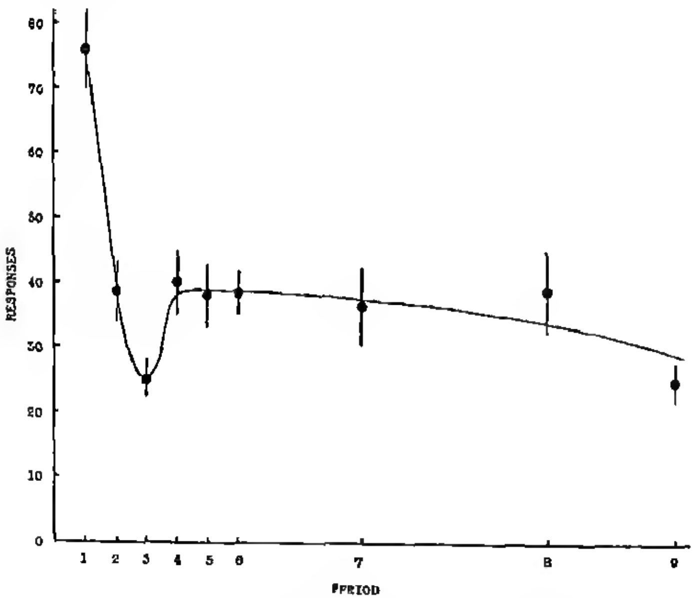

FIGURE 4

Mean Number of Non-Rewarded Responses Made by the Group of Forty Animals During the First Twenty Weeks of the Experiment

The bars represent the standard deviations of the means.

as a result of bar-pressing behavior in the experimental situation Eight responses during the first hour in the experimental environment may be taken, therefore, as the approximate level of random behavior The behavior of the experimental animals never dropped to this random level.

The tendency toward persistence of non-rewarded behavior is shown graphically in Figure 4 This curve represents performance during the first five months In plotting the curve responses occurring during the first hour only of each experimental period were considered The data for all animals were combined and a smoothed curve was drawn through the means for the entire group The vertical bars represent the standard errors of the mean points

Three distinct trends in the frequency of occurrence of non-rewarded responses are apparent Initially there was a rapid decrement in rate of response This decrement continued during the first three periods During the fourth period there was a marked increase in frequency of response The increase in rate of response between the third and fourth periods is more than three times its standard error Furthermore, the increase occurred in the case of all four groups. The change in the slope of the curve, therefore, may be considered highly reliable Following this abrupt increase in rate of response there was a third and longer period during which the animals continued to respond at a relatively constant rate, exhibiting a slow decrement which may have been due in part to increasing age

The final determinations were made on the 24th and 28th weeks All groups made fewer responses during these last two periods The animals which were allowed access to the running wheels made an average of only 13 responses an hour for the eight hours covered by the final tests During the previous eight hours they had made 28 responses an hour The animals which were not allowed access to the running wheels, on the other hand, averaged 19 responses an hour during the final periods as compared with 30 6 responses an hour made during the preceding eight hours These averages are based on animals upon which complete data were secured, exclusive of seven animals which died during the last two months of the investigation These findings indicate that the slow decrement in response continued and that the introduction of an alternative activity accentuated this decrement Even after seven months of experi-

mentation, however, the animals with access to the running wheels made more bar-pressing responses than were made by the control group the first time they were tested.

E. Analysis of Group Differences

The data secured during the first 20 weeks of the experiment reveal characteristic differences in the influence of the different combinations of work and deprivation. An analysis will be made of the interrelated influence of these two variables upon, first, the absolute number of responses made during each experimental period, and second, the relative change in rate of response from time to time.

1. *Group Differences in Number of Responses Made*

The mean number of responses and standard deviation of scores

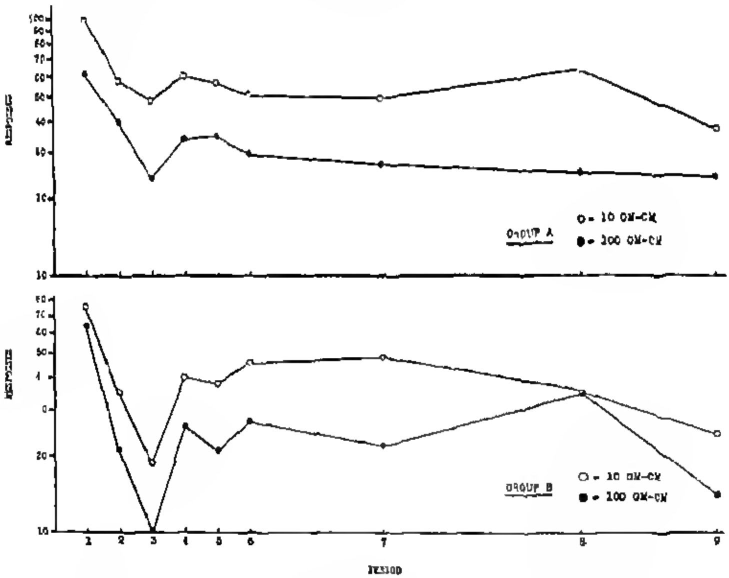

FIGURE 5

Means for Each of the Four Groups During the First Twenty Weeks of the Experiment

The *A* groups (upper) were at 48-hours of deprivation, the *B* groups (lower) were at 12-hours deprivation. Each point represents the mean number of responses made by a given group during the first hour of non-reward.

for each of the four groups during the first nine periods are given in Table 2 The same data are shown in the two graphs of Figure

TABLE 2
MEAN NUMBER OF NON-REWARDED RESPONSES MADE DURING EACH HOUR OF EACH EXPERIMENTAL PERIOD

Period		Group *A*-1 Mean	*SD*	Group *A*-2 Mean	*SD*	Group *B*-1 Mean	*SD*	Group *B*-2 Mean	*SD*
1		100	33 8	61	26 3	78	27.1	64	42 1
2		58	30 5	40	29 5	35	25 7	21	18 7
3		48	21 7	24	19 1	19	11 5	10	10 1
4		61	44 3	34	21 0	40	18 8	26	27 2
5		57	32 3	35	19 0	38	35 0	21	25 9
6		51	30 1	30	15 1	46	41 2	27	30 3
7	(1st hr)	49	53 8	27	19 0	48	42 0	22	22.9
	(2nd hr)	52	35 4	30	15 2	23	26 3	15	15 1
	(3rd hr)	32	21 8	29	18 7	20	18 5	15	17 2
	(4th hr)	22	13 8	15	7 1	11	10.5	6	8 2
8	(1st hr)	63	48 2	25	23 9	35	25 8	35	40 3
	(2nd hr)	70	28 9	15	9 5	43	31 2	25	20 8
	(3rd hr)	59	54 7	37	25 2	44	38 9	18	15 1
	(4th hr)	38	30 9	20	14 6	29	16 1	24	18 3
9	(1st hr)	37	23 8	24	21 4	24	10 6	14	8 6
	(2nd hr)	29	14 4	18	18 4	21	11 3	15	16 6
	(3rd hr)	27	19 3	7	6 0	32	27 7	17	17 9
	(4th hr)	16	10 1	14	14 7	28	30.8	15	13 5

5 In the upper graph is shown the data for the "*A*" groups, animals at 48 hours deprivation In the lower graph is shown the data for the "*B*" groups. The upper line in each graph is for the group doing only 10 gm-cm. of work Group *A*-1, it will be noted, exhibited the highest level of response These animals made the most responses during the first period and also on each of the subsequent periods. The rank of the groups, on the basis of number of responses made, was usually *A*-1, *B*-1, *A*-2 and *B*-2 The only departure from this exact order was during Periods 1, 3, and 8 These exceptions will be considered later. It can be concluded definitely from these results that animals which gave a larger number of spontaneous responses immediately after the withdrawal of food-reward also exhibited a comparatively higher rate of responding after many failures to secure food and many hours of non-reward This was true as long as the environment remained

constant Animals which were tested under conditions which led them to give few responses at first likewise gave fewer responses during later periods

These results do not agree with some of the interpretations drawn from similar studies Skinner has argued that groups of animals given equal training should eventually give the same number of non-rewarded responses "According to this view a certain number of responses, determined by the previous conditioning, is going to be observed in the extinction curve, no matter what the drive . Consequently, the lower the drive, the lower the initial rate in extinction, but the slower the decline in rate" (19, p 309) Thus Skinner would assume that the number of responses in reserve is fixed and remains unmodified by the conditions of elicitation Such a view finds no support in the present results The evidence indicates, instead, that the total number of responses which an animal will make under conditions of non-reward depends not only upon previous training and the length of time spent in the non-reward situation, but also upon the conditions of deprivation and work which prevail during such periods and likely upon other variables as well Animals at 48-hours of deprivation made more responses than animals at 12-hours, and animals doing 10 gm-cm of work made more responses than animals doing 100 gm-cm of work, regardless of what part of the total period of time is considered

2 *The Period of Rapid Decrement*

The rate of decrement during the first three periods was rapid The "*B*" groups showed a greater decrease in rate of response than the "*A*" groups, however Group *A*-1 made 100 responses during the first period and 48 during the third hour. Group *A*-2 dropped from 61 to 24 responses, Group *B*-1 from 78 to 19, and Group *B*-2 from 64 to 10 responses Critical ratios between these differences, computed by dividing the actual differences by the standard errors of such differences, are given in Table 3 It will be noted that the reliability of differences tends to increase from the first to the third period. Especially is this true for groups which were doing equal amounts of work per response, i.e., for differences between Groups *A*-1 and *B*-1, and between Groups *A*-2 and *B*-2

The semi-log plot in Figure 5 makes it possible to compare rate changes by comparing the slopes of the curves It will be noted at

TABLE 3

CRITICAL RATIOS OF THE DIFFERENCES BETWEEN GROUP MEANS FOR THE FIRST THREE PERIODS OF THE EXPERIMENT

	Group *A*-1 Period			Group *A*-2 Period			Group *B*-1 Period		
	1	2	3	1	2	3	1	2	3
Group *A*-2									
Period 1	2.7								
Period 2		1.3							
Period 3			2.5						
Group *B*-1									
Period 1	1.5			1.4					
Period 2		1.7			0.4				
Period 3			3.6			0.7			
Group *B*-2									
Period 1	2.0			0.2			0.9		
Period 2		3.1			1.6			1.3	
Period 3			4.7			1.9			1.8

once that the curves for the "*A*" groups are almost parallel, as are the curves for the "*B*" groups. An increase in work, in both instances, has displaced the curve downward without otherwise significantly modifying its characteristics. It would appear, therefore, that the condition of difficulty under which animals work neither accelerates nor retards the change in rate of response. This observation suggests that for purposes of further analysis of rate changes during this period we may combine the results of the "*A*" groups and also those of the "*B*" groups. In order to deal with rate changes alone, however, the data should first be freed from individual differences in absolute initial level of behavior. To effect this end the scores of each animal were multiplied by whatever constant was necessary to make the sum of the scores for three periods equal to

TABLE 4

MEAN NUMBER OF RESPONSES MADE BY THE "*A*" AND "*B*" ANIMALS DURING EACH OF THREE PERIODS WHEN THE TOTAL SCORES OF ALL ANIMALS ARE EQUATED

	Period 1	Period 2	Period 3	Total
"*A*" Groups	51	28	21	100
"*B*" Groups	66	21	13	100
Difference	—15	+ 7	+ 8	
Diff./S.D. Diff.	2.96	1.87	2.16	

one hundred. The means and standard deviations of these new scores for the *"A"* and *"B"* groups were then computed.

The significant data resulting from this treatment are given in Table 4. The mean animal in the *"A"* group made approximately two-fifths as many responses on the third day as on the first day, while the mean "B" animal made only one-fifth as many responses on the third day. This differential rate of decrement is reasonably reliable. The animals at a lesser condition of deprivation made relatively fewer responses than the more hungry animals during the first period, the difference being 2.9 times its standard error. On the third period they made relatively more responses and the difference, now in the opposite direction, was 2.2 times its standard error.

This differential rate of decrement during the early periods is in the direction postulated by Hull in an application of his goal-gradient hypothesis to some "field-force" problems. While the experimental situation is different, the present results are in agreement with the assumption "that a strong functional excitatory tendency is weakened after a given amount of frustration by a smaller proportion of its original strength than is a weak functional excitatory tendency" (12, p. 294).

A further interpretation of these differential rates of decrement can be made, also, in terms of the influence of work. It is obvious that after a long starvation period the capacity to do work is reduced. It has been demonstrated in the present study that when the task is made more difficult the animal responds less frequently. If it is now assumed that an increase in work-requirement has a greater effect upon starved animals than upon animals at lesser deprivation, then the present results can be explained on this basis.

Under the present experimental conditions the greatest number of responses, and therefore the greatest amount of work, occurred on the first period. During this initial period the *"A"* animals were not capable of doing as much work as the *"B"* animals. The number of responses made by the *"A"* animals, therefore, was in all probability limited to a greater extent by the work required in pressing the bar than was the number made by the *"B"* animals. Had work been less the *"A"* animals might have made many more responses during the first period, with no additional training or motivation. Two weeks later all animals were making fewer responses and doing less than half as much work. During this third period, con-

sequently, the activity of the starved animals was not limited by work as much as was formerly the case and it is not surprising that they showed a smaller reduction in rate of response. The rate of response, as experimentally determined on this group, showed a much smaller decrement than that of the "*B*" groups. Group *A*-2, for example, made fewer responses than either of the "*B*" groups during the first period, but on the third period made more responses than either of the less starved groups.

Several hypotheses can be formulated on the basis of such an analysis of the inter-related influence of deprivation and work. The work required in any situation will set a limit to the number of responses an animal will make in a restricted period of time, regardless of how highly motivated it may be. The greater the food-deprivation the more important will be the influence of work in setting a limit to the rate of performance. A peak of activity is to be expected early in the period of starvation in case (*a*) the response is one which involves a large energy expenditure, or (*b*) if training has been such as to lead to a very frequent repetition of the response. On the other hand, the peak of activity is to be expected only after a long starvation period in case (*a*) the task requires only a small expenditure of energy, or (*b*) if previous training has been such as to lead to infrequent responses.

3. *The Period of Recovery*

In all groups the rate of response increased during the fourth period. The animals did not again reach the level of activity exhibited during the first period, but the increase which did occur was statistically reliable for the 40 animals taken as a whole. The "*B*" groups, which had exhibited a more rapid decrement during the first three periods, made a greater recovery than did the "*A*" groups. This result agrees with the preceding analysis. The behavior of the "*B*" animals was not limited by work to such an extent as was that of the other animals, and therefore they should have shown a greater behavior change with any shift in motivation or change in the reinforcing state of affairs.

The tendency of behavior to become more stable as deprivation increases and more variable when animals are starved only a short time has been previously noted and reported. Elliott (7) rewarded rats after every choice, and found that variability in choice of path-

way was inversely related to hunger The ability to do work is likely an important factor in determining variability, whether variability be taken as the fluctuation in time required to eat uniform amounts of food, the number of different pathways traversed, or the change in rate of response from one day to the next Starved animals, for example, eat more slowly but also more continuously than animals with a greater food-reserve A change in the behavior of all the animals in the present experiment occurred during the fourth period, but the animals which reacted with a greater increase in rate of response were those more capable of doing an increased amount of work

One way to account for the recovery phenomena found in the investigation is to presume some change in the reinforcing state of affairs or a shift in motivation Such an idea is by no means new. It has been noted that organisms exhibit a tendency to continue activity in progress and to persevere in established behavior patterns Dashiell concludes that habituation may serve as a quasi-motive, and suggests that it "operates as a determining factor of such weight as almost to seem a force in itself" (6, p 125) The present data furnish support for this principle It would appear that in an environment where a single manipulative activity is possible it is only necessary to have the animal make the response a certain minimum number of times to insure that it will then continue to make responses regardless of whether or not it secures any extrinsic reward as a result of this behavior

Current theoretical and experimental work justifies the expectation that behavior established through the use of a reward may persist after removal of the reward Wendt (24) has presented evidence that an activity is inhibited only when some other reaction system takes its place Thus food-taking reactions would cease to occur only when other response systems natural to the situation gained dominance Brogden, Lipman, and Culler (4) report that unmotivated extinction remains incomplete More than the mere removal of shock was necessary to re-establish the equilibrium between alternative reflexes exactly as it had been before conditioning It would appear, therefore, that the perseverance of response found under the specific conditions of the present experiment is not a unique finding Instead it illustrates one of the basic principles covering the relation between motivation and behavior modification

The tendency for the rate of response to increase from its lowest level before becoming stabilized may or may not occur outside of the present experimental situation It would be of significance, also, to know whether or not the degree of training modifies the level of stable behavior eventually reached, or the extent of recovery

4 *Period of Stabilized Activity*

Six determinations of non-rewarded behavior were made between the 4th and the 20th weeks There were no marked changes in frequency of response during this period comparable to those noted during the early weeks Animals continued to respond at about the same rate during the first hour of each period During the longer four-hour experimental periods individual animals behaved differently Some responded less frequently toward the end of the four hours giving a negatively accelerated curve, while others gave approximately a straight line record The general trend of the records was from a negatively accelerated curve during the first few determinations of non-rewarded behavior to a straight line record during the latter periods This would indicate that the animals came more and more to respond at a uniform rate or in regular cycles during the entire time spent in the experimental situation

The greatest change in relative performance of different groups and also the greatest scatter of individual scores during this period of stabilized activity came on the 16th week This period was the one which was preceded by a break in the feeding schedule Even though periodic feeding was resumed, such a change in feeding might account for the variability of test scores on the following experimental period

The determinations made during the 20th week, constituting the 15th through the 18th hour of non-reward, showed the lowest rate of response since the end of the early period of rapid decrement. One group made approximately the same, one a smaller and two a slightly larger number of responses than each had made four months earlier during the third experimental period The decrement noted during this fifth month of the investigation may be attributed in part to a lessened activity level of the animals which were then over seven months of age

5. *Behavior Changes after the Introduction of the Running Wheel*

The number of cases in the sub-groups is too small to justify

analysis of the influence of the new activity upon each of the four experimental groups Some general conclusions can be drawn, however

Special interest attaches to the cyclic nature of the activity recorded by many of the animals which were given access to the running wheels. In Figure 6 is shown three four-hour records secured during

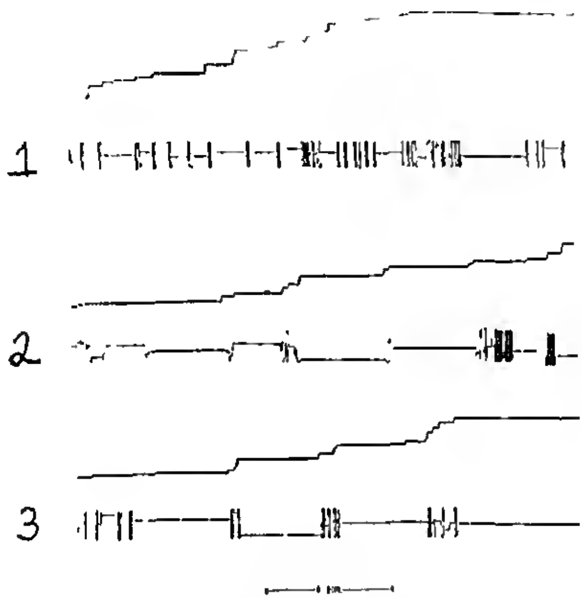

FIGURE 6
THREE FOUR-HOUR RECORDS OF BAR-PRESSING (UPPER TRACING) AND SPONTANEOUS RUNNING ACTIVITY (LOWER TRACING)

this time The top line in each record is the bar-pressing response curve, and the lower line records rotations of the running wheel The second and third records were made by animal No 20 during the 24th and 28th weeks At both times this animal was doing 100 gm-cm of work at 12 hours of deprivation Each record shows clearly that the animal alternated between short periods of activity and longer periods of inactivity During the activity phase, furthermore, animal No 20 engaged in both running and bar-pressing activity It would run for a brief period, then leave the running

wheel and make several bar-pressing responses, then run again, etc This sort of behavior was characteristic of many of the animals Their records suggest the periodic nature of contractions of the muscles of an empty stomach and could be presumed to show a relation between periods of stomach contractions and periods of activity There was no check, in the present investigation, on whether or not stomach contractions were actually occurring during these periods of activity, but other investigators have demonstrated such a correspondence Richter (16) found that periods of activity in the rat occurred at intervals of from one to two hours, which is roughly the periodicity of the activity recorded in Records 2 and 3.

The top record in Figure 6 was made by animal No 27 during the 24th week This animal was studied while doing 100 gm-cm of work at 48-hours of deprivation. It failed to show any of the cyclic activity characterizing the other records.

Animals in the *"A"* groups showed a greater reduction in bar-pressing with the introduction of the new activity than did the *"B"* animals This difference has low statistical reliability, but it is in the direction of previous results and of previous interpretations The *"A"* animals were stimulated to run and thus to expend a greater amount of energy than formerly, and this should have left them less capable of carrying on the old activity. The *"B"* animals, on the other hand, had been without food for only 12 hours, and should have been able to run and at the same time continue the old activity with only a slight decrease in rate

The fact that animals persisted in the bar-pressing habit, even though it was possible for them to engage in an activity normally more closely associated with hunger, i.e., running, furnishes final evidence for the permanence with which the bar-pressing activity had been established through repeated elicitation.

F General Discussion and Evaluation

Studies of animal motivation have been very extensive and have combined many degrees of training, work-requirement, drive, reward, and punishment Each of these variables undoubtedly plays a part in determining the frequency of occurrence of any given behavior sequence It is not surprising, in view of the complexity of factors involved, that apparent inconsistencies have been noted in comparing the results of different studies, since in many cases only one variable

has been controlled in the same manner in any two experiments Many of these studies of motivation have been concerned in one way or another with the question of when activity reaches a peak in the course of starvation Here, especially, divergent results have been reported The hunger drive has been reported as maximal at anywhere from a few hours to six days of starvation

Gross bodily activity has been found to reach a peak at six and again at 21 hours (13), and also at 48 to 72 hours (16) Animals crossed an electrical grill more frequently after 48 to 72 hours of starvation (23), but were more likely to cross another similar obstruction after 144 hours (15). Most investigators have found in the case of white rats that the rate of eating decreases after a 24-hour period of starvation (2, 14) Fitts (10) found that this was true when feeding pellets of solid food, but failed to find such a reduction in eating after 72 hours of starvation when feeding powdered food. Heron and Skinner (11) gave their rats small pellets of food every four minutes and found that the rate of bar-pressing increased up to the 96th hour of starvation

As Elliott and Bousfield have pointed out, "An examination of such varied behavioral effects, each of which seems to follow its own individual law, suggests that there are a number of distinct mechanisms operating in food deprivation" (9, p 94) The present results lend support to such a view and indicate the nature of some of the mechanisms involved Behavior changes cannot be attributed to the influence of any one variable alone, and no single behavior measure gives an accurate picture of the strength of the drive

In order to evaluate the conditions of motivation existing under different experimental conditions not only must the external manifestations of the drive condition be determined, but this behavior should be analyzed further in relation to the cost to the organism and the difficulty of the behavior under each experimental condition It should be borne in mind at the same time that frequently there is a tendency toward perseveration and fixation of response. An organism cannot be expected to work beyond its capacity, but on the other hand it does not follow that the mere removal of an external reward will lead to a complete loss of any given behavior pattern Such a complete loss did not occur in the present experiment Any theoretical consideration of the stimulus-response relationship should take this phenomena of perseveration in established behavior into consideration as well as the limiting effect of excessive work

G. Summary

1. Forty white rats were trained to push down a small bar in order to get food. They were given 30 reinforcements in making this response. For the next seven months periodic determinations were made of the rate at which the animals continued to make non-rewarded responses. In order to analyze the inter-related influence of various factors in producing changes in this behavior the rats were divided into four groups and each group was studied at a different combination of deprivation and work.

2. Three stages of non-rewarded behavior were noted: (*a*) a period of rapid decrement during the first few weeks, (*b*) a period of partial recovery in rate of response, and (*c*) a much longer period of stabilized behavior during which the animals continued to respond with little change in rate from week to week.

3. Animals doing a small amount of work made more responses at all times than did animals which were required to do a greater amount of work.

4. Animals at 48-hours of deprivation with minor exceptions made more responses than did animals at 12-hours of deprivation.

5. Animals at 12-hours of deprivation exhibited a very rapid decrement in rate of responding during the first stage of non-rewarded activity and following this a marked recovery. The animals at 48-hours of deprivation, on the other hand, showed more stability of behavior. This was attributed to the fact that the lessened energy reserve of the starved animals reduced their capacity for doing work and hence rendered them less likely to respond with marked behavior changes to variations in the reinforcing state of affairs.

6. The introduction of a running wheel as an alternative activity during the 6th and 7th months was followed by a reduction in rate of bar-pressing. Even under these conditions, however, the rate of response was not as low as the original level of a control group which had received no training reinforcements. During the experimental periods animals tended to engage regularly in periods of activity during which they would run and also make bar-pressing responses. Between such cycles of activity were longer periods of inactivity.

7. The data furnish additional evidence that when the reward is removed at the end of a training period the animal does not always cease to make the response established by this training.

References

1 Bousfield, W A. Certain quantitative aspects of the food-behavior of cats *J Gen Psychol*, 1933, **8**, 446-454

2 Bousfield, W A., & Elliott, M H The effect of fasting on the eating-behavior of rats. *J. Genet Psychol*, 1934, **45**, 227-237

3 ———— The experimental control of the hunger drive *J Genet Psychol*, 1936, **15**, 327-334.

4. Brogden, W J., Lipman, E A, & Culler, E The rôle of incentive in conditioning and extinction *Amer J Psychol*, 1938, **51**, 109-117

5. Bruce, R. H The effect of lessening the drive upon performance of white rats in a maze *J Comp. Psychol*, 1938, **26**, 225-248

6. Dashiell, J F. Fundamentals of General Psychology. Boston Houghton Mifflin, 1937 Pp xxxii+655.

7 Elliott, M H The effect of hunger on variability of performance *Amer J Psychol*, 1934, **46**, 107-112

8 ————. Drive and the characteristics of driven behavior *Psychol Rev.*, 1935, **42**, 205-213

9 Elliott, M H, & Bousfield, W A Two basic mechanisms in motivation *Psychol Rev.*, 1936, **43**, 94-99

10. Fitts, P. M Experimental studies of hunger-motivated behavior and learning. Thesis, University of Rochester, 1938

11 Heron, W T, & Skinner, B. F. Changes in hunger during starvation. *Psychol. Rec*, 1937, **1**, 51-60.

12 Hull, C L. The goal-gradient hypothesis applied to some "field-force" problems in the behavior of young children *Psychol Rev*, 1938, **45**, 271-299

13, Ligon, E. M. A comparative study of certain incentives in the learning of the white rat *Comp Psychol Monog.*, 1929, **6**, No. 2, 1-95.

14. Maclagan, N F The rôle of appetite in the control of body weight. *J Physiol*, 1937, **90**, 385-394

15 Moss, F. A A study of animal drives *J Exper Psychol.*, 1924, **7**, 165-185.

16 Richter, C. P. Animal behavior and internal drives *Quart Rev Biol*, 1927, **2**, 307-343

17 Skinner, B. F Drive and reflex strength I & II. *J Gen Psychol*, 1932, **6**, 22-48.

18 ———— On the rate of formation of a conditioned reflex *J. Gen Psychol*, 1933, **7**, 274-286

19. ———— Conditioning and extinction and their relation to drive *J Gen Psychol.*, 1936, **14**, 296-317.

20. ———— The Behavior of Organisms New York Appleton-Century, 1938 Pp 457

21. Stone, C P A study of perseverance and frustration in working rats Paper given before the 45th meeting of the A P A, 1937.

22. Wada, T An experimental study of hunger in its relation to activity *Arch. of Psychol*, 1922, **8**, No 57, 1-65

23. Warden, C J Animal Motivation: Experimental Studies on the Albino Rat New York. Columbia Univ Press, 1931 Pp xii+502

24 WENDT, G. R An interpretation of inhibition of conditioned reflexes as competition between reaction systems *Psychol Rev*, 1936, **43**, 258-281.

25 WILLIAMS, S B Resistance to extinction as a function of the number of reinforcements *J Exper Psychol*, 1938, **23**, 506-522

26 YOUNG, P T. Preferences and demands of the white rat for food *J Comp Psychol.*, 1938, **26**, 545-588

27 YOUTZ, R E P Reinforcement, extinction, and spontaneous recovery in a non-Pavlovian reaction *J Exper Psychol*, 1938, **22**, 305-318

Department of Psychology
University of Tennessee
Knoxville, Tennessee

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1940, 57, 193-198

AN ENIGMA OF THE HEREDITY-ENVIRONMENT QUESTION*

Otterbein College

John E Wenrick

Hereditarians and environmentalists have long argued over the problem of the relative importance of heredity and environment as the two factors that determine the origin and development of intelligence Partisans for each of these positions have written a voluminous literature in support of their particular point of view From the standpoint of the numerous assumptions involved in an inquiry of this type, it is possible that experimental evidence might be found which tends to support either—or/and—both sides of the question The experiments themselves may be unquestionable, but the interpretations derived therefrom suggest the necessity of a critical evaluation of some of the fundamental issues

It is the purpose of this paper to examine critically some of the assumptions underlying the construction of intelligence tests and in the light of this critique determine the validity of the experimental results relative to the hereditary-environmental question. In spite of the excellent work of Newman (7) and others who have definitely shown us the pitfalls into which one might fall while investigating this and similar problems, and the extreme care with which the conclusions must be drawn from the data, the problem of the validity and reliability of the experiments is vital, at least in so far as it concerns or determines the amount of space given this topic in textbooks, where whole sections have been devoted to a consideration of this problem To quote from Ruch (8, p 180), "The sum of our hereditary influences is certainly greater than that of measurable

*Received in the Editorial Office on May 8, 1939

environmental influences" Similar statements are abundant in current literature, and some are even more specific in suggesting the fixed ratio of these influences.

The claims or findings of the investigators in this field have, as we have pointed out, been at odds as to which is the most important factor in the determination of intelligence If the hypothesis that neither is more important than the other be true or, if either one or the other be true, it would be necessary to make various changes in educational and social procedures Naturally, if one is more important than the other then it is necessary that we find out how much and in what direction growth and development lie Numerous studies have suggested evidence that heredity is more important than environment in the determination of intelligence In the light of the data that have been secured in the investigations pertaining to this issue, it is pertinent to raise certain questions as to the validity of our tools and the techniques that are employed in the investigation of this problem. The techniques that have been employed have been summarized by Newman (7), Schwesinger (9), and others and need not be repeated in detail here. Basically, they are techniques employing co-twin control. These studies have been concerned with fraternal and identical twins reared in similar and dissimilar environments.

More specifically, the points that we are interested in examining here, are with reference to the assumptions underlying the construction and development of intelligence tests, and the possibility that these assumptions themselves may determine or influence the data that may be secured in an investigation of this kind. In most cases the tools that have been employed as the criteria or as the source for measuring the relative effects of environment and heredity are the various intelligence and performance tests, or more specifically, the *IQ* rating on these tests. In order to see if the intelligence tests themselves are valid instruments or tools for studying this question, we must examine briefly the assumptions underlying the construction of a test of intelligence. Out of the great number of assumptions underlying the construction of intelligence tests, three are pertinent to this problem.

(*a*) The test must be relatively new. It must avoid as much as possible anything that is learned by the subjects under specific conditions, e g., factual material learned in the school, such as historical

dates or facts of sociology, etc In other words, the items must be based upon those circumstances and events that are unconsciously learned (acquired) in the general environmental situation. Factors of specific knowledge or training must be eliminated Those items are to be incorporated which are based upon the commonality of environmental situations *This may be interpreted in one way at least as the ability to adapt to relatively new situations.* The question as to the difference between factual knowledge and intelligence as two entirely different concepts needs no elaboration

(*b*) Differential items must be selected that will pick out successive age levels, that is, the items at the different age levels must be *selective* for that age group There must be some things that a twelve year old child can do that a nine year old child cannot do The method of selection of these items, then, is *empirical* and is based upon the average or norm for that particular age group in question. The important point to consider here is that these items must be *differentially selective,* that is they will identify a child's mentality as belonging to a particular age group.

(*c*) The reliability of the test, in some measure at least, follows out of the considerations outlined in number two If our test is to be reliable we must have items at the different age levels that are consistent in identifying a particular mental age development or at least identifying them to the extent that they fall within a given age norm This means *empirically* that those items or test are stricken out which have a great scatter or variability and according to Terman and Merrill (12), age arrangement is set up which will make the mean mental age identical with the chronological age in order to obtain an *IQ* that is as close as possible to 100 The reliability of our test then will be some function of the amount of scatter or variability within the individual items comprising the test If Assumption 3 follows out of Assumption 2, then it also follows that factors underlying Assumption 3 will in part determine the results obtained under Assumption 2.

In view of these factors which underlie the construction of intelligence tests and the empirical method at work in the selection of various test items, we might ask the following question · *Are internal consistency of the test and the empirical method of selection of the various mental age items arguments that the environmental factors have been eliminated or at most held to a minimum?* This refers

specifically to the factors outlined in Assumption 1, that the test items must not test the factual knowledge but must probe the unconsciously learned behavior. This may be interpreted to mean the ability to adapt to new situations as well as the degree of commonality of environment among the subjects in question I interpret this to signify the hereditary factor. In other words, the internal consistency of the test, both from the standpoint of the reliability of the items and the *empirical method of selection* of the items at the particular age levels may be taken as the indicator that the item in question is not susceptible to specific training This conclusion seems valid in light of some of the experiments relative to the influence of coaching or training upon raising the *IQ* and those investigations concerned with the constancy of the *IQ* over a long period of time Grave's (3) and Minogue's (6) investigations support the idea that the *IQ* rating is relatively constant in spite of the many influences which seemingly might raise or lower the *IQ* The studies concerning the differences in intelligence levels between rural and urban populations, Jones, Conrad, and Blanchard (4), and the study of Sherman and Henry's *The Hollow Folk* (11), have demonstrated clearly the difficulties involved in the empirical method of standardization of intelligence tests. These results support the thesis that variability in performance levels or intelligence levels is not a function of specific but of the general environmental influences. Heredity and environment go hand in hand in determining behavioral patterns, and the empirical method of standardization tends to eliminate environmental factors as such in the determination of intelligence This will justify the statement that the "items are selective with regard to the common hereditary factors," that is, Assumptions 2 and 3 may be taken as evidence for the elimination of the varying differences in ability that the environmental influences produce [1]

We may come to one conclusion therefore: the tests are prejudiced If these three assumptions eliminate environmental factors as the predominant determiners of differences in intelligence, then it must follow that our intelligence tests are necessarily constructed to bring to light those factors and conditions that are primarily determined

[1]This leads one into a consideration of the complex problem of maturation and we do not mean to say that intelligence is not influenced by environment, but that the factors involving the construction of the tests of intelligence eliminate the factors of environment as determiners of variability

by the biological and hereditary development of the individual If such be the case, the very tools which we use to study the relative effects of heredity vs. environment are themselves already heavily-loaded in favor of the so-called hereditary factor Is it any wonder that the majority of studies produce evidence or results that heredity is more important than environment?

We have eliminated those factors which may be influenced by environmental conditions and then use the very tools from which these environmental conditions have been eliminated to investigate the very problem we set out to solve From the standpoint of the construction of our intelligence tests we have eliminated to a large degree one of the factors which we wish to study

Newman (7) has pointed out the fallacies that may arise from investigations of this type when too little of the evidence has been presented, and it may be inferred that he was arguing along some such line as this paper has outlined. The basis for this inference is the following quotation from the book, *Twins: A Study of Heredity and Environment*

> It is apparent from several of the comparisons made that the relative effect of hereditary and environmental differences is also a function of the type of trait Any fixed ratio of these two factors for all traits and conditions is thus impossible. We must consider their relation always in connection with the kind of trait and grade of environmental difference

Some traits are more susceptible to environmental factors than are others If twins are reared together most of the differences may be due to a nature factor, but if the twins are reared under strikingly different environments the nature factor will have a relatively greater influence We would expect to find, then, such factors as personality and performance skills as being more affected by environmental changes than those found in intelligence test studies This again may be taken as evidence that the environmental factor has been eliminated or at least minimized in the selection of those items that go to make up our intelligence tests.

In the light of these considerations, the stage has been set a priori in favor of the finding of the greater influence of heredity as the determining factor in the development of intelligence If this be true, there are two conclusions that we may draw from these studies,

which employ the intelligence tests as tools for the measurement of differences between heredity and environment First, they do not represent a valid picture of the rôle of heredity as a determining factor of intelligence. Second, at most they are not experimental demonstrations of the relative influence of heredity and environment, but are *arguments for the validity and reliability of the tests.*

Summary

If our intelligence tests are based upon the factors which measure the conditions that are relatively uninfluenced by environmental changes, then these tests when used to investigate the relative effects of the factors as determiners of intelligence are an argument for the validity and reliability of the intelligence tests and nothing more. This is not itself an unimportant finding, but let us not draw more from the experiments than an examination of the fundamental postulates will warrant.

References

1 BURKS, B The Relative Influence of Nature and Nurture upon Mental Development *27th y'r'b'k Nat Soc. Stud Educ*, 1928, Part I, 219-316.

2 CATTELL, P Constant changes in the Stanford-Binet *IQ*. *J. Educ. Psychol.*, 1931, 22, 544-550

3 GRAVES, K. The influence of specialized training on tests of general intelligence *Teach. Coll, Columbia Univ Contrib Educ*, 1924, No 143

4 JONES, H E, CONRAD, H S, & BLANCHARD, M. B. Environmental handicaps in mental test performance *Univ. Calif Pub Psychol*, 1932, 80-83.

5. LEAHY, A M. Nature-nurture and intelligence *Genet. Psychol Monog*, 1935, 17, 236-308.

6. MINOGUE, B M The constancy of the *IQ* of mental defectives *Ment. Hyg*, 1926, 10, 751-758.

7 NEWMAN, H H, FREEMAN, F U., & HOLZINGER, K Twins A Study of Heredity and Environment Chicago: Univ. Chicago Press, 1937. (P 349)

8. RUCH, F L Psychology and Life. Chicago Scott, Foresman, 1937

9. SCHWESINGER, G C Heredity and Environment. New York Macmillan, 1933 (P 313.)

10. SHAFFER, L F. The Psychology of Adjustment New York. Houghton, Mifflin, 1936

11. SHERMAN, M., & HENRY, T K. The Hollow Folk. New York. Crowell, 1933.

12 TERMAN, L M, & MERRILL, M A Measuring Intelligence New York. Houghton Mifflin, 1937.

Otterbein College
Westerville, Ohio

The Journal of Genetic Psychology, 1940, 57, 199-205

KINETIC ASPECTS OF GROWTH: A GRAPHIC STUDY*

Nevada City, California

CLARENCE QUINAN, M D

This article is based upon the quantitative graphic time data obtained in experiments made with 200 unselected normal persons Its purpose primarily is to describe a remarkable growth phenomenon; one which stabilizes itself definitely at the age of puberty Incidentally, however, chiefly in order to show that they differ from one another significantly with reference to a time frame, attention is called also to a number of other cognate growth processes.

PRELIMINARY CONSIDERATIONS

In the dictionary sense growth is said to be "the gradual increase of a living organism by natural process." Usually a growth process appears to be regarded as being an accretion phenomenon; one which completes its full course within a limited period whether that period be less than a year as in the case of a plant, or three score years and ten as in the case of a human being

But there seems to be also another, non-cyclical, ever-expanding, evolutionary type of growth Apparently it is representative of a form of energy which is irresistible, ubiquitous, cumulative in its action and of such a nature that even inanimate things exhibit unmistakably indelible marks of its activity. That this is so becomes obvious, for example, on comparing two such objects as a paleolithic weapon and the modern cyclotron, or a Roman chariot and a China-clipper Indeed, according to the doctrine of substance advanced by Leibnitz (2), this latter form of growth energy seems to press ever-forward in a universal field of action; to determine all natural forms as well as the onward progress of civilization.

In the case of a human being, rather coarse growth movements become distinguishable at a fairly early period; and the remarkable

*Accepted for publication by Harold E Jones of the Editorial Board, and received in the Editorial Office on May 15, 1939

thing about them is that each one seems to pursue an independent time schedule of its own

Thus the fetus in utero (*a*) kicks as early as the fourth month; and (*b*) during the fifth month its heart sounds become audible "The fetus in utero also sucks its fingers," (1) thus exhibits (*c*) inchoate phases of mental energy. Moreover, the oncoming citizen of this world (*d*) adds about two inches to the length of its body in the course of each month and (*e*) during a like period makes a gain of about 12 ounces in weight.

Hence the fetus displays (*a*) palpable movements of its extremities, the force concerned being undoubtedly that which subsequently expresses itself as the graphic speed factor, (*b*) rhythmic heart beats —so many per minute, (*c*) early traces of mental function, and, (*d*) three mass factors these are capable of expression definitely in terms of length, bulk, and weight respectively

Even in utero therefore the respective activities of mind and matter seem to express themselves simultaneously and progressively in the form of the developing embryo.

Now from Locke's standpoint (3), in regard to the factor which concerns the "fundamental verity" or "bottoming" of an organic growth process some pregnant bits of concrete objective evidence are available These have to do with the nature of the productive force and with some of its specific activities.

Jacques Loeb (4) states emphatically "that animal movements depend upon light in the same way as the movements of plants." For example, respecting the large marine Annelid, *Spirographis spallanzanii*, he goes on to say:

> If the light falls upon the animal from one side only, heliotropic curvatures make their appearance in the tube. The animal turns its oral pole toward the source of light and bends its tube until the axis of its radially expanded gills lies in the direction of the rays of light

In other words, radiant energy of solar origin produces motion in this creature. But light also produces growth A considerable body of experimental evidence supports the theory which holds that vegetable chloroplastids function as an energy-transforming mechanism, that this mechanism converts the kinetic energy of certain permanent gases into the potential energy of solids—e g , into that

of a carbohydrate Since this conversion process goes on incessantly in the vegetable world growth is its inevitable corollary. Whence it follows that light, motion, and growth are merely forms of energy And naturally there is also an entropy equivalent

Living substance, the ultimate reality, represents, then, organized energy integrated by photosynthesis Indispensable seems to be the sun's light Other desiderata are controlled heat, moisture, and the presence of suitable electrolytes and catalyzers Thus chlorophyll reappears every spring, vegetation leaps upward.

Incidentally, respecting the intra-uterine movements already mentioned, one special point here should be stressed namely, that each one of these movements seemingly must sustain a relationship of some sort to that partition mechanism upon which the body's symmetry depends For presumably a suitable distributing agency must exist the purpose of which is to divide growth energy into equal halves—right and left

It may be added, finally, that in the case of a human being the immensely complicated energy transformations of which a growth process is representative always take place at an optimal temperature close to 36 5 *C.*

In the writer's opinion, the foregoing evidence suffices at least to suggest the idea that life and motion are closely related energy states Hence, since energy is indestructible, may it not perhaps be inferred that life too is imperishable? That is to say, in the individual, restricted, non-evolutionary sense may it not constitute simply a cycle of solar energy, precisely as in the case of a plant?

Conceivably, such a cycle might be supposed to consist of, (*a*) a developmental period—i e , extending from the instant at which the primitive centrosome divides to that at which the body attains full stature, (*b*) a middle period of maturity; and, (*c*) a terminal period of fructification or of productivity on the part of the mind

Consider now the kinetic aspects of growth as manifested at successive age levels. Seemingly an infant is never still unless it be asleep It clutches things, kicks, stretches, yawns, and has no sphincter control Later on in life, during the school age, young people romp, contend with one another in games of every description, quarrel and fight; energy, obviously a redundant commodity, is expended lavishly. The children vary greatly as to age and as to size. Some are graceful, many are awkward, all are noisy Look

at one more scene. On a university campus full grown men and women move about quietly. They have passed through the developmental stages of growth, exhibit degrees of endocrine dignity seemingly because, in both sexes, the organs of reproduction are become functionally active.

Evidently then, as regards the visible expenditures of kinetic energy, a striking decrescendo movement characterizes that period which terminates at the age of puberty. Since the progress of this energy peculiarity can readily be studied by means of quantitative time data referable to successive annual age levels, some data of that sort will now be introduced.

Material and Technic

The working material was obtained from 200 unselected normal persons—100 orphan girls and 100 University of California Undergraduates. The girls were tested in the San Francisco Roman Catholic Orphanage, an institution which offers comprehensive educational facilities within its own walls. Ten groups of the inmates were studied, the grades represented being those from the second year of grammar school to the third year of high school. The age range was from eight years to 17 years. The average age of the university men was nearly 19 years.

As regards technic, the procedure followed was quite simple. No preparatory instructions were issued further than to request the subjects to write in their customary style. Each person tested was timed with a stop-watch on this sentence: "He protested that he couldn't eat the twenty-two tarts, and the truth is that he ate twenty." In this way a quantitative time value was obtained referable to a constant unit of graphic work. The learning factor was entirely eliminated as regards this particular sentence since but a single experiment was made in each instance. The model object placed before the subjects was typed, double-spaced, in such a fashion as to form a neat three-line block. All the girls wrote with the same fountain-pen; the men used their own pens. Each reactor copied down the test words on a library card 127 mm. by 76 mm. in size.

Results

The 100 university men (Table 1) ran up a group average value of 37.8 seconds on the model sentence. The extreme figures noted

TABLE 1

Graphic Time Data Obtained by Means of a Stop-Watch and a Model Sentence from 100 University Undergraduates Having a Mean Age of 19 Years

	Frequency distributions (Seconds) 26-29	30-39	40-49	50-59	Total	Mean speed value sec
Men, No.	6	66	19	9	100	37.8

were 59 seconds and 26 seconds respectively. However, it will be observed that 91 per cent of these subjects made records of 49 seconds or under. It should be stated that 65 per cent of the men performed the work within 38 seconds while eight turned off their specimens in 30 seconds or less.

On the other hand (Table 2) the mean values found respectively as regards the 10 groups of girls tested ranged from 253 seconds at eight years of age to 46.6 seconds for girls 17 years old. The respective extreme values obtained were 305 seconds and 38 seconds,

TABLE 2

Graphic Time Data Obtained by Means of a Stop-Watch and a Model Sentence, at Successive Age Levels, from 100 Orphan Girls Ranging in Age from Eight Years to 17 Years

Children Age	No.	Frequency distributions (Seconds) 38-49	50-99	100-149	150-199	200-249	250-305	Mean velocity sec.
8	5				1		4	253.0
9	10			5	3	2		168.4
10	10		1	6	2	1		134.7
11	15		6	8	1			105.3
12	14		11	3				84.0
13	15	1	14					59.0
14	12	1	10	1				68.3
15	8	2	6					57.0
16	8	3	5					56.6
17	3	3						46.6
	100	10	53	23	7	3	4	

the larger figure being referable to a girl aged eight years, the smaller, to a girl 16 years old

As for the trend of the graphic speed factor toward a stabilization level, the figures in Table 2 speak for themselves. It will be noted that what seems to be a critical change in the velocity rate occurs at the age of 13 years.

Comment

Apparently in any given instance it is the purpose of a normal growth process to complete in full detail a perfect structural pattern Thus nature spins out exceedingly fine ramifications for each Purkinje's cell, causes the uppermost branches of the evergreen tree to adjust themselves to the figure of a true cone Naturally elaborations such as these connote a large time factor. But are not all dynamic phenomena time functions?

Of course the ability to produce handwriting is an accomplishment. Certainly it is not an inherent vital endowment Consequently all graphic data are essentially adventitious in character and, probably for that reason, are but little used as tools of research. Nevertheless, quantitative graphic time data in one respect appear to be unique and invaluable, for by such means alone is it possible to study step by step as it were the kinetic aspects of growth during the developmental period

Summary

1 A dynamic growth phenomenon characteristic of the period of adolescence is described

2. By means of a stop-watch and a model sentence time data were secured from 200 persons—100 men having a mean age of 19 years and 100 orphan girls from eight years to 17 years of age

3. The mean time values obtained were for the men, 37 8 seconds, for girls 17 years old, 46 6 seconds, for eight year old girls 253 seconds.

4 Seemingly the age level at which the graphic speed value stabilizes itself is that which marks also the functional maturity both of the central nervous system and of the reproductive organs

References

1 De Lee, J B The Principles and Practice of Obstetrics Philadelphia, Saunders, 1915

2 Leibnitz, G *Encyclopedia Brit*, 1911

3 Locke, J An Essay Concerning the Understanding and a Treatise on the Conduct of the Understanding Philadelphia Kay & Troutman, 1849

4 Loeb, J The Heliotropism of Animals and Its Identity with the Heliotropism of Plants Chicago Univ Chicago Press, 1905

133 Grove Street
Nevada City, California

The Journal of Genetic Psychology, 1940, 57, 207-210

AESTHETIC MATURITY*

Lane Technical High School, Chicago, Illinois

JAMES M GUNTHORP

In order to learn the effect of age on aesthetic appreciation a comparison was made between the aesthetic judgments of adults and of children of various ages To begin with 22 pictures were obtained from the Fine Arts Department of the University These were then paired In some pairs the pictures were of nearly equal merit, as in the case of two landscape scenes by Corot; in other pairs the two pictures differed widely in this respect, for example *The Gleaners* by Millet and *Hunting the Deer* by Couse In addition three pairs were made from pictures taken from a calendar because it was thought these would appeal to boys and girls in the elementary grades The 14 pairs were shown to a group of adults interested in art, a club of professional men and women, and the preference for one of the pictures in each pair was noted in the case of each individual. After 50 days the test was repeated to determine the constancy of preference for each person The judgments of these adults were then used as a norm by which to judge the preferences of junior groups

To facilitate this comparison the degree of preference of the adults for each picture was determined and each given a rating accordingly Some pictures were preferred by more adults than others Each picture which was preferred by a majority of adults was given as a rating the difference between the number of adults preferring it and the number preferring its rival. This rating represents the excess vote or degree of preference for the picture, which appears to be a proper basis for rating pictures grouped in pairs. The preferred picture in each pair was designated "*A*" and its rival "*B*"

The pictures were shown in the public schools of Vermillion, South Dakota, and to a group of students in the University of South Dakota In two elementary schools, designated "*A*" and "*B*," Grades 2, 4, and 6 were used for the test, and in the high school one class each in Grades 8, 10, and 12 and the art department

*Received in the Editorial Office on June 6, 1939

The groups varied in size, ranging from 13 to 27 persons in each. Each pupil's choices were noted at two different times with an interval of 50 days between as in the case of adults

After each showing of the pictures each member of a junior group was given a score consisting of the sum of the ratings of all the "A" pictures chosen by him. For example, one child chose only two pictures in agreement with the adults, each of these having a rating of four Her score was, therefore, eight. The child making the highest score chose all but two pictures in agreement with adults. Her score was 96, the sum of the ratings of her 12 "A" pictures The average of the two scores given each child was used in making the comparisons. Table 1 shows average scores by age and sex. Average scores of girls exceeded those of boys

TABLE 1
AVERAGE SCORES BY AGE AND SEX GROUPS

Age	Score: Boys	Score: Girls
7	36	44
8	38	45
9	34	43
10	47	39
11	39	46
12	51	48
13	52	58
14	43	52
15	49	64
16	57	61
17	52	68
18	59	75

Of interest is the fact that increase in average score accompanied advance in age, the coefficient of correlation being 53 when the group was taken as a whole, and 43 and .41 for boys and girls, respectively, when the group was divided on the basis of sex.

The relation between art appreciation and cultural background is not so readily observed Three different methods of determining the cultural level of the pupils were used

(*a*). School attainment, with reference to both the study of art and general school level, was used as a criterion of culture A marked advance in agreement with adults accompanied higher school level attainment, the one exception being the art class composed of girls from Grades 9, 10, and 12 who had a higher average score than Grade 12

(*b*) A second method of estimating cultural level used was the judgment of teachers concerning the home from which each pupil came. Three or four teachers of each class were asked to give the relative cultural level of the homes represented in it (the comparison to be on the basis of income, house occupied, pictures on the walls, recreational advantages, clothing of the children, parents' education, and whatever else seemed pertinent). The ratings of the three or more teachers forming judgments of each home were averaged to give it a composite rating. It is recognized that not too much importance is attachable to these ratings as there are intangible factors in the culture of a home which are not open to observation.

However, this approach to the study yielded some possibly noteworthy results. In six of the nine groups of boys those above the median or middle score were rated by their teachers as enjoying a higher than average home cultural level. The same was true of seven of the ten groups of girls. Seventeen correlations were made of cultural level ratings and scores taken by age groups (whole year). In nine of the 17 groups there were coefficients of correlation above .4; the others were scattered between .21 and —.48, three of them being negative. A correlation of cultural level ratings with score, taking all cases in one group, gave a coefficient of .10. As mentioned above, in evaluating the significance of this correlation it should be recognized that the cultural rating given a child by one group of teachers probably does not mean the same as an identical rating given another child by a different group of teachers. By the use of the multiple correlation formula,

$$r_{12.3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{1 - r^2_{13}}\ \sqrt{1 - r^2_{23}}}$$

the correlation between cultural level and score uninfluenced by age differences was found to be .20.

(*c*) Another check of the influence of cultural level was suggested by the contrast between the two elementary schools. School *"A"* serves part of the city having comfortable living conditions, School *"B"* serves a very low income section. In one of the three classes tested in each school, School *"A"* excelled School *"B"* in agreement with the adults. In the other two classes, the reverse was true. Apparently the lower standard of living did not lead to a poorer appreciation. This finding is believed to be of interest, but no explanation of it is offered.

The relation of consistency of judgment was another phase of the study undertaken. It was found that the average adult chose 11 pictures consistently on the two showings, but this was equalled by three other groups: the art class, the girls of age 16 and the boys of age 13. The youngest girls, aged seven, excelled 13 other groups in consistency. The coefficient of correlation between consistency of preference and age of boys was found to be .01±.06. The corresponding coefficient for girls was the same. The coefficient between consistency and score was —.17±.06 for boys, —.07±.08 for girls. Correlations of consistency and cultural level also gave negligible coefficients.

Summary and Conclusion

So far as the tests used may warrant, these conclusions may be drawn:

1. There is a gradual approach to maturity of aesthetic judgment from early ages to adulthood.

2. Girls are somewhat more mature than boys.

3. (*a*) A noticeable relation obtains between training in art and art appreciation and a possible relation between general school level and appreciation; which, however, may be attributable to age as well as culture. (*b*) Some relationship appears between aesthetic judgment and the cultural level of pupils (as rated by teachers). (*c*) There is no positive relation between standard of living and appreciation.

4. No correlation was found between consistency and either age or score or cultural level.

212 Franklin Street
Barrington, Illinois

The Journal of Genetic Psychology, 1940, 57, 211-217

THE INFLUENCE OF MONTH OF BIRTH ON THE INTELLIGENCE OF COLLEGE FRESHMEN*

Department of Psychology, University of Pittsburgh

OMAR C HELD

According to the astrologer the date of one's birth determines everything that he is or will become Those who are born under Venus shall be very gay lovers, pleasant and delicious, and most commonly they shall have black eyes and little brows; while those who are born under Saturn shall be false, envious, and full of debate and full of law Those who are born under Mercury shall be very subtle of wit, and shall be a devout person to God, and have good conscience, and shall be crafty in many sciences.

What scientific validity if any do such pronouncements have? Professor Huntington (3) in his book *Seasons of Birth* concludes that the season at which people are born has far greater importance than is generally supposed in determining health, competence, or success.

Professor Huntington, in studying the relationship between time of birth and competence, used as his criterion of success the fact that the person was listed in some book of eminent people Professor Huntington as a result of his study concludes that everything seems to favor the late-winter and early-spring babies

There have been a number of other studies in this field where intelligence test scores were used as a measure of capacity Blonsky (1) in a study of 453 children tested by the Binet Test found seasonal differences in the mean *IQ* The mean *IQ* is highest for spring and lowest for winter However, the differences are slight Table 1 summarizes Blonsky's findings

The mean *IQ* is highest for spring and lowest for winter Slight as these differences are, and without calculating the significance of the differences, Blonsky concludes that the month of birth influences mental development.

Pintner (4) made use of group intelligence test results on 4925 school children of all grades in his study. He concluded that the

*Received in the Editorial Office on July 22, 1939

TABLE 1
MEAN INTELLIGENCE QUOTIENTS OF CHILDREN IN THE VARIOUS SEASONS ACCORDING TO BLONSKY

Season	Mean *IQ*	*N*
Spring	84 3	115
Summer	81 5	111
Autumn	81 3	114
Winter	80 1	113
Total	82 1	453

month of birth did not influence the *IQ* of the cases studied. The winter or cold months had means slightly below the warm months, but none of the differences were statistically reliable

Pintner and Forlano (5) tabulated *IQ*'s for 17,502 children in all grades up to the end of high school They found a reliable difference between the mean *IQ* for the winter months and the mean *IQ*'s for the other seasons The differences between the other three seasons were negligible.

Fialkin and Beckman (2) tabulated the sigma ratings of 5717 men clients of the Adjustment Service of New York City who took the Pressey *Senior Classification and Verification Tests* The ratings used in this study were expressed in positive terms by setting zero at minus 2 5 sigmas and 10 at plus 2.5 sigmas, the intervals between these extremes being expressed as the number of ½ sigmas

Fialkin and Beckman conclude from their study that month of birth is a factor in intelligence. They say that in general, those born in months of moderate temperatures are superior to those born in cold months However, the difference between the means for spring and winter was only 1 80 times its standard error, which is not usually considered indicative of statistical reliability.

In the present study the psychological percentile scores were tabulated for 2327 University of Pittsburgh students. The test used was the American Council Psychological Examination which is given to freshmen after they are admitted to the University. The study included both men and women students

A distribution table by month of birth was made Table 2 shows the mean psychological percentile rank by months

The lowest mean rank was shown by college students born in April, the highest by those born in September. Figure 1 shows graphically the monthly means

TABLE 2
MEAN PSYCHOLOGICAL PERCENTILE BY MONTH OF BIRTH

Month	Mean Psych percentile	Number of cases
January	48.6	179
February	49.9	194
March	49.5	197
April	47.3	215
May	50.6	201
June	51.8	195
July	50.6	194
August	48.1	207
September	52.8	188
October	49.2	201
November	50.3	184
December	50.6	172

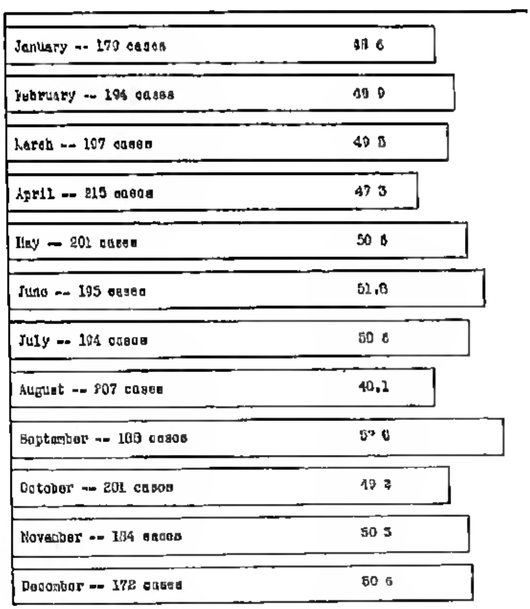

FIGURE 1
MEAN PSYCHOLOGICAL PERCENTILE BY MONTH OF BIRTH

TABLE 3

RANK OF MONTHS IN FOUR STUDIES OF THE RELATION OF MONTH TO MENTAL RATING*

	Studies			
Month	Pintner	Pintner-Forlano	Fialkin-Beckman	Present study
January	11	11	7	10
February	10	12	12	7
March	9	9	9	8
April	6.5	3	5	12
May	5	6	2	3.5
June	3	1.5	3	2
July	4	5	4	3.5
August	8	8	10	11
September	2	1.5	1	1
October	1	7	8	9
November	12	4	11	6
December	6.5	10	6	3.5

*Means are in terms of *IQ* in Blonsky, Pintner, and Pintner-Forlano studies; in terms of half sigmas in Fialkin-Beckman study, and in terms of percentile ranks in the present study.

In Table 3 the months were ranked in order of decreasing means so as to compare the ranking of the months with that in other studies.

Reference to Table 3 shows that September is the highest ranking month in three of the four studies. In the Pintner-Forlano study September shares highest ranking with June. In the present study April was the month with the lowest mean. The difference between the means for September and April is only 1.88 times its standard error. This is not considered a statistically reliable difference

Table 4 shows the means for the four seasons for the present study and various other studies

In each of the four previous studies the same general trend exists. In each instance there was a difference particularly between spring and winter. Blonsky did not determine the statistical reliability of the difference in his study Pintner in his 1931 study found that the mean for spring exceeded that for winter, but the difference was not statistically reliable Pintner and Forlano in the 1933 study found the mean for spring exceeded that for winter and the difference was statistically reliable Fialkin and Beckman in their study found that the difference between spring and winter was only 1.80 times its standard error, which is not generally considered a statistically reliable difference

In the present study summer had the highest mean and winter the

TABLE 4

MEANS OF MENTAL RATINGS ACCORDING TO SEASON OF BIRTH*

Season	Blonsky	Pintner	Pintner-Forlano	Fialkin-Beckman	Present study Mean	Present study No
Spring Apr.-June	84 3	97 20	102 35	6 69	49 8	611
Summer July-Sept	81 5	97 20	102 05	6 66	50.4	589
Autumn Oct.-Dec	81 3	97 10	101 85	6 58	50 0	557
Winter Jan.-Mar.	80 1	95 95	100 65	6 53	49 3	570

*Means are in terms of *IQ* in Blonsky, Pintner, and Pintner-Forlano studies, in terms of half sigmas in Fialkin-Beckman study, and in terms of percentile ranks in the present study

lowest The differences between any of the seasons was very slight The difference between summer and winter was only .65 times its standard error, which is not statistically reliable

Another comparison made in previous studies has been between the cold months and the moderate months. December, January, February, and March were selected as the cold months and April, May, September, and October as the moderate months Table 5 summarizes this data for this study and for previous studies.

TABLE 5

MEAN MENTAL RATINGS FOR THOSE BORN IN MODERATE AND IN COLD MONTHS ACCORDING TO FOUR STUDIES

Research	Moderate months Number	Moderate months Mean	Cold months Number	Cold months Mean	Ratio of difference to its standard error
Pintner		97 61		96 11	2 37
Pintner-Forlano		102 29		100 79	4 06
Fialkin-Beckman		6 68		6 55	1 72
Present Study	802	49 6	742	49 9	0 20

*Means are in terms of *IQ* points in the Pintner studies, half sigmas in the Fialkin-Beckman study, and of percentile rank in the present study

Reference to Table 5 shows that only in the Pintner-Forlano study the ratio of the difference to the standard error is high enough to constitute a statistically reliable difference. In the present study there is no essential difference between the means.

In Table 6 the means for the warm months, May to October, are compared with the means for the cold months, November to April

TABLE 6

MEAN MENTAL RATINGS OF THOSE BORN IN WARM AND IN COLD MONTHS ACCORDING TO THREE STUDIES

Season or period of birth	Pintner	Means Pintner-Forlano	Present study
Warm May-October	97 55	102 15	50 5
Cold November-April	96 15	101 30	49 3

The cold months in each instance show a lower mean, but the differences are very slight and in neither of the three studies was it statistically reliable In the present study, based upon 1186 births in the warm and 1141 in the cold months, the ratio of the difference to the standard error was only 1 03.

The following conclusions seem warranted from the present study:

1. The month of birth has no influence upon the intelligence of the college students who were the subjects of this study.

2. The season of birth has no influence upon the intelligence of this group of college subjects who were studied

3 There is no statistically reliable difference between the intelligence of the subjects of this study born in months of moderate temperature and those born in cold months.

4 There is no statistically reliable difference between the intelligence of the subjects of this study who were born in warm months and those who were born in cold months.

5. This study found those born in September to have the highest mean mental rating, and in that respect is in agreement with the Fialkin-Beckman study and also with the Pintner-Forlano study where September shared first place with June.

REFERENCES

1 BLONSKY, P P Fruh-und Spatjahrkinder Jahrbuch fur Kinderheilkunde, Leipzig, 1929.

2 FIALKIN, H N, & BECKMAN, R O. The influence of month of birth on the intelligence test scores of adults *J Genet Psychol*, 1938, 52, 203-209.

3 HUNTINGTON, E Season of Birth New York Wiley, 1937
4 PINTNER, R Intelligence and month of birth *J Appl Psychol*, 1931, **15**, 149-154
5 PINTNER, R, & FORLANO, G The influence of month of birth on intelligence quotients. *J Educ Psychol*, 1933, **24**, 561 584

Department of Psychology
University of Pittsburgh
Pittsburgh, Pennsylvania

APPARATUS

The Journal of Genetic Psychology, 1940, 57, 219-220

A SLIDING FRAME FOR USE IN REGISTERING CHOICE*

Cambridge, Massachusetts

Marie L H Forbes

The drawing (Figure 1) shows a frame sliding on horizontal rods over a wooden platform on which a sheet of paper with punches

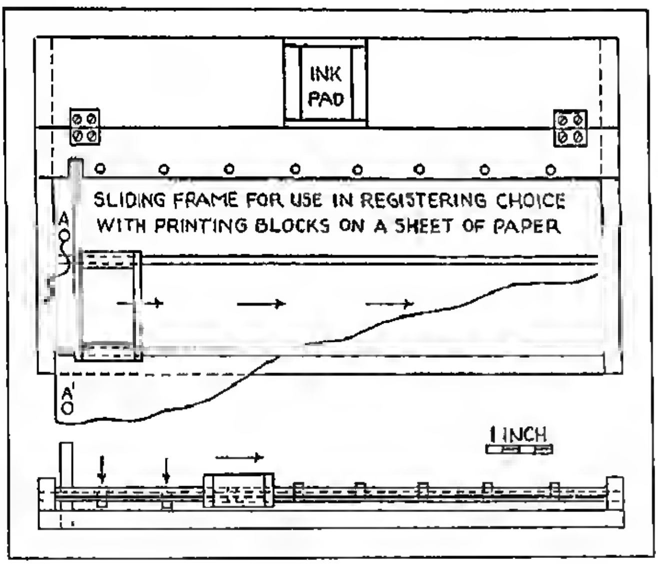

FIGURE 1

*Received in the Editorial Office on June 10, 1939

(*A*, *A'*) in the left margin is pushed between cleats away from the operator and held in position by a peg.

The writer gratefully acknowledges the examination of the materials by Dr. Walter F Dearborn, Director of the Psycho-Educational Clinic, Harvard University

8½ Ash Street Place
Cambridge, Massachusetts

The Journal of Genetic Psychology, 1940, 57, 221-246

BOOKS

The *Journal of Genetic Psychology*, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than $2 per printed page and not more than $3 per printed page

Conditions Only those books that are listed below in this section are eligible for such reviews In general, any book so listed contains one or more of the following traits (*a*) Makes an important theoretical contribution, (*b*) consists largely of original experimental research, (*c*) has a creative or revolutionary influence in some special field or the entire field of psychology, (*d*) presents important techniques

The books are listed approximately in order of receipt, and cover a period of not more than three years A reviewer must possess the Ph.D degree or its equal in training and experience

Procedure If among the books listed below there is one that seems important to you, you are invited to write a review of that book It is not necessary to make arrangements with the Editor Just send in your review It does not matter if the book in question has been reviewed before

(1938)

STRONG, E K Psychological Aspects of Business New York McGraw-Hill, 1938 Pp 629

BOAS, F The Mind of Primitive Man (Rev ed) New York Macmillan, 1938 Pp 285

SHERMAN, M Mental Conflicts and Personality New York Longmans, 1938 Pp 319

GUTHRIE, E R The Psychology of Human Conflict New York: Harper, 1938 Pp 408

MURPHY, G, & LIKERT, R Public Opinion and the Individual New York Harper, 1938 Pp. 316

KATZ, D, & SCHANCK, R L Social Psychology New York Wiley, 1938 Pp 700

ELLIS, W D A Source Book of Gestalt Psychology New York Harcourt, Brace, 1938. Pp 403

WOODWORTH, R S Experimental Psychology New York Holt, 1938 Pp 889

SEASHORE, C E. Psychology of Music New York McGraw-Hill, 1938 Pp 408.

SKINNER, B. F The Behavior of Organisms New York Appleton-Century, 1938 Pp 457

TERMAN, L M Psychological Factors in Marital Happiness New York McGraw-Hill, 1938 Pp 474

MURRAY, H A, *et al* Explorations in Personality New York Oxford Univ Press, 1938 Pp 761

LANDIS, C., & PAGE, J D. Modern Society and Mental Disease New York Farrar & Rinehart, 1938 Pp. 190
MUNN, N L Psychological Development New York: Houghton, Mifflin, 1938 Pp 582
RASHEVSKY, N Mathematical Biophysics Chicago: Univ Chicago Press, 1938 Pp 340.
KOHLER, W The Place of Value in a World of Facts New York Liveright, 1938 Pp 418
OGDEN, R M The Psychology of Art New York Scribner, 1938 Pp 291

(1939)

ROSETT, J The Mechanism of Thought, Imagery, and Hallucination New York: Columbia Univ Press, 1939. Pp. 289
PRATT, C C The Logic of Modern Psychology. New York. Macmillan, 1939 Pp 185
DOLLARD, J, DOOB, L W., MILLER, N. E, MOWRER, O H, SEARS, R. R., *et al* Frustration and Aggression New Haven: Yale Univ Press, 1939 Pp 209
GESELL, A., CASTNER, B M., THOMPSON, H, & AMATRUDA, C S. Biographies of Child Development New York Harper, 1939 Pp. 328
GOLDSTEIN, K The Organism. New York World Book, 1939. Pp 533
BARTLETT, F C, GINSBERG, M, LINDGREN, E. J, & THOULESS, R H. The Study of Society New York Macmillan, 1939 Pp 498
MAIER, N R F Studies of Abnormal Behavior in the Rat. New York Harper, 1939 Pp 81.
WECHSLER, D The Measurement of Adult Intelligence Baltimore: Williams & Wilkins, 1939 Pp 229.
WOODWORTH, R S Psychological Issues. New York Columbia Univ Press, 1939. Pp 421
KARDINER, A The Individual and His Society New York. Columbia Univ Press, 1939 Pp 503
STEVENS, B The Psychology of Physics. Manchester, England: Sherratt & Hughes, 1939. Pp 282.
ELLIS, H. My Life: Autobiography New York: Houghton-Mifflin, 1939

1940

CANTRIL, H The Invasion from Mars A Study in the Psychology of Panic Princeton Princeton Univ. Press, 1940. Pp. 228
RASHEVSKY, N Advances and Applications of Mathematical Biology Chicago Univ Chicago Press, 1940 Pp 214
BURLINGAME, L L Heredity and Social Problems New York McGraw-Hill, 1940 Pp 369.
DENNIS, W The Hopi Child New York D. Appleton-Century, 1940 Pp 204
HILGARD, E R, & MARQUIS, D G Conditioning and Learning New York Appleton-Century, 1940 Pp 429
HARTMAN, G. W, NEWCOMB, T, *Editors* Industrial Conflict A Psychological Interpretation New York Cordon, 1940 Pp 583
DOOB, L W The Plans of Men New Haven Yale Univ Press, 1940 Pp. 411.
HULL, C. L, HOVLAND, C I, ROSS, R. T, HALL, M, PERKINS, D T, & FITCH, F B Mathematico-Deductive Theory of Rôte Learning New Haven: Yale Univ Press, 1940 Pp 329

CRITICAL REVIEWS OF RECENT BOOKS

(*Burrow, T. The Biology of Human Conflict. New York: Macmillan, 1937. Pp. 435.*)

REVIEWED BY WILLIAM GALT

Formulations within the domain of science invariably set out with certain assumptions which are accepted as fact. Subsequent investigations are then oriented in relation to these assumptions. With the early astronomers such an accepted fact or *prima materia* was the assumption that our world constituted the centre of the universe and that the sun and lesser planets revolved around it. For many centuries astronomical investigations were made from this geocentric background. But although many precise observations were recorded and much ingenuity exercised in building up a framework in support of the assumed premise, no cosmologically true relationships could be deduced, precisely because of the unsoundness of the fundamental assumption.

Such has been the universal history of scientific thought. It only rarely occurs in any scientific field that the scientist turns about and questions the accepted assumptions or the "given facts" from which he has previously set out upon his researches. It is a rare occasion when an investigator challenges the basis of his investigations and, instead of keeping to the well-beaten path that leads *outward* from his assumptions, faces about and follows an *inward* course to the examination and evaluation of the assumptions themselves. Wherever such a course has been pursued there has been opened the possibility of unprecedented progress in respect to the field in question.

Copernicus adopted such a course when he made bold to question the earth's central position in the astronomical system—when he dared to substitute the heliocentric theory of the universe for the geocentric conception of Ptolemaic tradition. Modern medicine likewise followed this inward path of inquiry when it had the temerity to challenge the time-honored tradition that illness and death were the result of black magic, an "evil spirit" or the anger of God, and rejected cures "performed" through the superstitious rites of exorcism, incantation, and prayer.

In the field we have been pleased to call the humanities, or in the field that has to do with "man's relation to man"—in the sphere of

man's mental and emotional aberrations, his crimes and punishments, his domestic and international competition and conflict, his family relationships, his social institutions and his religious ceremonials—it is very specially true that there have grown up among us certain thought-habituations which have acquired the position of fundamental assumptions and *from which* investigations of human behavior have invariably set out

The significance of the researches of Dr Trigant Burrow, of which he has recently given a systematic account in his latest book, *The Biology of Human Conflict*, is precisely that in the field of human behavior and its disorders he has reversed the ordinary procedure Instead of setting out from a background of unquestioned assumptions in regard to human nature and the motivations to behavior, Burrow has chosen the far more difficult task of directing his inquiry *inward* upon himself and his social community and investigating the validity of such behavioral assumptions as are the heritage of our social structure and culture

Obviously, this is not an easy task Traditional assumptions in any field are not to be questioned and re-evaluated with impunity The idea of a mobile earth outraged the deep-rooted prepossessions of the 16th century just as the conception of the bacterial origin of disease-processes did violence to the assumptions of the century just past. Furthermore, the offense is increased a thousand-fold where the tenets in question involve human behavior and its intricate emotional involvements The impetus toward traditional interpretations in the field of man's own motivations and behavior is not a condition which can be *externally* observed and corrected, it is a condition *internal* to the processes of the observer himself

But if the investigations of Burrow and of his experimental group have been arduous and time-consuming, they have brought fresh conceptions of fundamental importance to the field of human behavior In the following pages I shall mention some of these altered viewpoints and try to indicate their far-reaching implications

Early in his researches Burrow realized that if he was to make headway in his study of the basic factors and values in human social behavior it was necessary that the private prejudices, the emotional biases and immaturities, in short, the "personal equation" of the research student (and of the normal community generally) be "corrected for" Taking his cue from the objective sciences where a

consensus of observers is employed to rule out any inconsistency of data resulting from impaired sense organs or faulty observations, an experimental group was formed for the study within itself of the fundamental attributes of human behavior and of social interrelations The story of this experimental group is a long one and it would take us too far afield to enter upon it here Perhaps I may say, however, that the observations which grew out of this concerted endeavor did not represent a mere majority of individual conceptions, ideas, opinions, or prejudices Rather, after years of endeavor and analysis, there was a synthesizing effect in which the group as a behavioral unit was able to reach below the outer crust of habituations and prejudices which commonly color the personal reactions of the individual It was able to reach a level of motivation at which the organism sees and interprets clearly phenomena placed before it even where these phenomena have to do with the cherished biases of man's own behavior

In the opinion of the reviewer the most fundamental of the conceptions Burrow has introduced is his thesis regarding the basic cohesion or principle of coordination operative in man as a species His position is that the species, man, represents fundamentally a phylo-organismic unit or whole and that the behavior of the individual primarily derives its motivation and control from this intrinsic matrix of organismic solidarity In this view, the primary motivation to man's behavior has its origin within the species and is always oriented in respect to species-solidarity The "socialization" of the individual, however conventionally valuable, is only a superficial and transient veneer, when compared with man's phylobiological basis of unity and coordination From this organismic frame of reference the individual as a separate unit of motivation represents a mere conceptual abstraction The social group and not the individual man is the primary behavioral unit This intrinsic social organization constitutes for man an indissoluble part of his make-up It is the primary and ineradicable basis which convention cannot set aside So that the researches of Burrow in the field of human behavior abrogate an egocentric conviction in regard to the origin and meaning of human behavior and posit a phylo-centric basis of interpretation just as the astronomical investigations of an earlier epoch substituted a heliocentric for a geocentric conception of the universe

But as far as concerns man's behavior, this is not the whole account. It seems that somewhere in the age-long history of man's biological descent there has occurred a developmental *faux pas*. For the group-analytic investigations of Burrow make evident that the coöperation and interdependence characteristic of biological foundations is not operative in such sophisticated human groups as the family, the church, political parties, or nations.[1] These groups operate from secondary, ideological, and individualistic motives of behavior and not from motivations basic to the species man as a whole. Although individuals and groups may be very sentimental in their efforts to help others, this very "helpfulness" stems from a feeling of difference. It does not arise from a sense of man's primary solidarity as a phylum. Such benevolences are in no sense a guarantee against the most ruthless and socially disjunctive behavior. Oscar Wilde must have had something of this sort in mind when he remarked that "charity causes a multitude of sins."

Burrow has brought this anomaly in man's behavior into line with biological data generally by his discovery through group experimentation that man as a racial organism is suffering from a "social neurosis".[2] It should be emphasized that the term "neurosis" is not used figuratively by Burrow. He holds that all individuals on the present "normal" basis of adaptation are, from a biological viewpoint, maladjusted and disordered in their behavior-processes. As I have already indicated, this position does not represent an a priori view, but has resulted from a painstaking and exhaustive analysis of social groups in experimentally controlled situations continuing over a period of many years.

In this analysis it was found that the social reactions of man which are habitually thought to be normal are, from an unbiased observation, as "partitive" and as lacking in total biological health as the isolated behavior of the neurotic patient. It was found that the normal individual, for all his socially acceptable modes of interchange

---

[1]Instances of this biological principle permeating animal life are given in profusion in the books of Professor W. C. Allee. *Animal Aggregations. A study in General Sociology*, Chicago: Univ. Chicago Press, 1931. Pp. 362. *Animal Life and Social Growth*. Baltimore: Williams & Wilkins, 1932. Pp. xii + 159.

[2]This concept of the social nature of the neurosis was first developed by Burrow in a paper published in 1914. "The psychanalyst and the community." *J. Amer. Med. Asso.*, 1914, 62, 1876-1878.

and communication, in reality makes no more intrinsic contact with his fellows than does the neurotic individual who is constantly preoccupied with his private phantasies, wishes, and dreams The normal, like the neurotic, is constantly thinking and acting in terms of his fancifully individualized self For man has substituted for the organism's primarily motivated behavior a purely "social image" of behavior The individual has established an image of himself as an isolated behavior-center entitled to private wishes, desires, and motives that bear no relation whatever to the total motivation of man as an organismic whole This *social image* of himself and of his behavior now dominates the entire field of man's interrelational processes. The normal no less than the neurotic lives in an unreal social world of which he is the central arbiter

Man, however, is still an animal He has not permanently lost his biological heritage of species-solidarity, inter-individual coordination, and social integrity However overtly individuated, the element or individual can again be brought under the sovereignty of the primary action-centre that motivates the phylum It is in this connection that Burrow's observations on what he has called "the preconscious mode" are of special interest. For the preconscious mode represents a continuation into adult life of the unitary phase of behavior that characterizes the infant in its early socio-biological relationship to the maternal organism Burrow has emphasized that the mother does not at first represent for the infant an external object or personality

> I came to see that the mother was at first not by any means the infant's love-*object*, but rather, if I may so express it, its love-*subject* Lacking as yet the faculty of symbolic projection that is to be acquired only with later training, the infant organism does not, at first, look out toward or feel itself in objective relation to the mother The child and the maternal organism are coterminous There is no sense of separation, of distinction, of otherness For the infant there is at first no interval, no distance, which must be bridged in order to reach the object opposite it, outside itself As yet there exists no object, no opposite, no other self, nor aught outside itself

In adult life the intimations of this preconscious mode—"a mode of completion and fulfillment, of uninterrupted confluence and totality"—are found chiefly in dreams, in poetry and in music These

intimations occur in moments when one is wholly absorbed in music, totally engrossed in a beautiful sunset or "gazing over a limitless expanse of moon-lit sea."[8] This mood is characterized by self-effacement, by tranquillity, and by a lack of competitive concern. It would seem to represent a symptomatic expression of the biological continuity inherent in man as a species. The preconscious mode indicates that the principle of species-solidarity has not been lost in the dim history of man's past, but that it is an active principle present in each generation until it is perverted afresh to partitive, individualistic concern and self-seeking.

As a result of the undue individuation that has occurred among the elements or individuals of the species and that has brought about the subjective conviction in man of an egocentric, inter-personal basis of behavior, there has arisen an extremely arbitrary and vacillating measure for judging behavioral expressions. Our daily experiences are redolent of these inconsistencies. For example, my judgment in regard to a trespass on my own part will differ greatly from my judgment of the same trespass when committed by some one else. Again, I incline to interpret the act of a friend very differently from the act of an enemy. The identical act may be approved in the one case and condemned in the other.

But this is a commonplace in so far, at least, as our projective experience is concerned. Burrow goes farther. He emphasizes that our customary standards of good and bad, of right and wrong, have become divorced entirely from the criteria of biological fitness or unfitness which governs the behavior of other animals. To quote:

> The originally total, organic sense of wholeness, coordination and fitness, or the basic rightness that is common and consistent throughout all organisms of the species, has been shifted to a mere social image of wholeness or rightness. This shift has been accompanied by a phenomenon of behavior whereby the isolated, separated individual has been vested with a pseudo-authoritative or proprietary "right." Right is *his* right. It is *his* private possession.

And Burrow goes on to develop the invigorating, if unorthodox, thesis that every man at all times, whether he be gangster or mission-

[8]"Happiness," a brief essay by James Norman Hall that appeared in *The Atlantic Monthly*, November, 1939, is undoubtedly driving at something of this sort.

ary, does what he "feels at the moment to be right" In other words, the individual has so aligned his sense of right with his sense of personal advantage or private gain that "right" and "self-advantage" have become one and the same

I do not know anywhere a keener analysis of this insidious moral dichotomy in the social motivation of man than is presented in Burrow's chapter on *The Social Fallacy of Right and Wrong* Here he shows how generally we accept as perfectly consistent the fallacious situation presented daily in our courts of law He cites the common instance in which two lawyers of equal standing and integrity take opposing sides in relation to a prisoner being tried for homicide One attempts to sway the jury to the belief that the defendant is guilty of the deed and should be hanged, the other brings to bear all of his oratory and ingenuity to persuade the jury that the man is innocent and should be acquitted And more tellingly still, this investigator shows how a competent psychiatrist in the pay of the State will find a prisoner sane whereas a second psychiatrist of equal eminence and training, but deriving his fee from the defense, will pronounce this same individual insane and hence not responsible for his acts The author very sagely concludes that if, in their own conduct, behavior-experts are guided by a sense of right that is completely unstabilized because wholly inspired by their sense of personal gain, they are in no different case from the criminal whose sense of right is coterminous with his personal gain Consistent with his phylobiological frame of reference it is precisely Burrow's position that the disturbed balance in basic values inseparable from man's dichotomous sense of right and wrong is endemic throughout human behavior on its present plane of adaptation

This brings us to another important point in Burrow's thesis If it is true that in civilized, adult communities one always acts in accordance with one's sense of personal gain, and if this sense of personal gain has become identical with one's "sense or sensation of right," it follows that a moral, symbolic, or linguistic approach to behavioral conflicts and inconsistencies cannot be effective It follows that the sense of gain that is socially applauded and called "good" may be biologically as unfitting and distorted as the sense of gain that is socially outlawed and decried as "bad" Burrow points out that the only way we can gain access to this baffling problem of human behavior—one's own and other peoples'—is to discard our

mental approach to it and reach down into the immediate physiological sense of our own behavioral inconsistency, that is, into the conflicting physiological patterns underlying our own behavior-motivations, individual and social.

Proceeding along these lines, Burrow has made substantial headway in observing and correlating the physiological alterations concomitant to disordered behavior He has gathered evidence to show that, coincident with a patient's division in feeling and thought, there has occurred a division within his physiological reaction-patterns This division and delimitation of response is shown externally in the various symptomatic manifestations with which we are all familiar —in conflict, indecision, anxiety, tenseness of response, stammering, ritualistic behavior, etc. This lack of unity and coordination in essential behavior-systems is not, however, characteristic alone of the neurotic personality The author's investigations show that where the reaction-average of our ordinary social interrelationships—the responses we fondly call "normal"—is submitted to an inclusive analysis, there is unquestioned evidence of imbalance and distortion within physiological patterns motivating the behavior of the community as a whole

These phylobiological investigations have centred specifically upon the process of attention as it relates the organism to its environment They have differentiated two broadly contrasting types of attentional reaction and have established the physiological correlates of these response-patterns The first is the usual pattern of attention with which we are all familiar. It is essentially projective in mechanism and is primarily associated with the use of the symbol or of language Where this form of attention is bound up with inter-individual interests and relationships, it has been found that its projective evaluations are in large measure colored by moralistic connotations, that they are biased by anxious self-concern and other egoistic behavior-trends which are competitive and socially disaffective Dr Burrow has referred to this type of the organism's response as "partitive," and recent experiments show that it is characterized physiologically by small, rapid eye-movements, a rapid respiratory rate, and general muscular hypertension

The contrasting attentional mode which Burrow calls "integral attention" or "cotention" represents, on the contrary, the organism's immediate response as a whole to the environmental situation This

type of response is not projective It is not primarily mediated through the symbol or through verbal mentation This mode of attention is not occupied with externalized images and inter-individual affects It is concomitant to physiologically recognizable tensions within the cephalic segment precisely when these customary images and affects are held in abeyance. Subjectively, cotention is accompanied by an absence of mental conflict, of petty self-concern and individual competitiveness. The organism experiences a sense of greater poise, of a larger reservoir of interest and motivation, and is in more complete rapport with its physical and social environment Physiologically, cotention is characterized by a marked diminution in number of eye-movements and a greatly reduced rate of respiration

Since the publication of Dr Burrow's book there has been the instrumental recording of modifications in other physiological systems coincident with the induction of cotention These alterations will be fully described in a report which Dr Burrow is now preparing in accordance with the statement at the close of his book

There are other interesting formulations in *The Biology of Human Conflict* which I shall not attempt to cover in this review Specific chapters in the book will make their special appeal to students in different fields of inquiry In Chapters II, VI, and XIII the psychoanalysts will find a brief consideration of the psychological systems of Freud, Jung, and Adler, and the development of a new concept in regard to the important theme of sexual repression Burrow emphasizes that in the neurotic individual it is not the true biological drive of sex (an organismic manifestation) that is repressed, but the individual's arbitrary and phantastic social images of sex The animal psychologist should find material of special interest to him in Chapters VI, VIII, and XI where there are discussions of the work of Pavlov and his associates, and the linking up of the phenomenon of the "experimental neurosis," as it is induced in animals, with neurotic manifestations occurring in man

The physiologist and the neuro-anatomist will find material of greater interest to their field in the discussion of the organism's total and partitive reaction-systems, and of the rôle of the symbol and of language in partializing man's bionomic and social adaptation (Chapters V, VI, and IX) In this connection Burrow draws interesting analogies with Coghill's fundamental neuro-anatomical studies The anthropologist and the sociologist will find stimulation in the

radically new conception of the original nature of man's consciousness and of socialized institutions, while the educator will come upon numerous implications which may, after careful study, be carried over to his particular sphere of interest Finally, the biologist should be especially intrigued by this study of *human conflict* because of the broad biological background and frame of reference which orient Dr Burrow's approach to the problem of human behavior and its disorders.

The criticism has often been made that Dr Burrow's terminology is difficult and his meaning not always easy to grasp There will undoubtedly be those who will feel there is ground for this criticism in the present volume I should like to remind the reader, however, that the basis from which Burrow sets out entails a fundamentally altered bionomic orientation It has been necessary for him not only to fashion new tools for investigating behavior but also words, or verbal tools, for conveying his conceptions and findings After all, it is easy enough to say something, and say it well, when one is only saying what has been said thousands of times before It is a different matter where one is handling new concepts and necessarily requires new instruments with which to handle them In view of this circumstance, it should not seem surprising that these verbal instruments are as yet somewhat awkward and not as expert and facile as they will undoubtedly tend to become in future years It is the earnest hope of the reviewer that in this process of refinement the terminology will not lose in pragmatic vitality and meaning Language, unlike any other form of tool, often suffers through the process of being made facile and easy of operation A formulation fraught with deep significance can in the course of this process become a glib formula conveying no intrinsic feeling or meaning Words and phrases in their refinement have an only too frequent tendency to become mere verbalization divorced from the organism as a whole and the intrinsic meaning with which it has imbued them

In conclusion, may I say that Dr Burrow's thesis is not one that may be approached with timid, hesitant steps We live in a moment of significant revolutions in all fields Our very civilization stands at the cross-roads Man can not proceed further along the old, accustomed ways He is at a definite impasse New outlooks, new foundations, new internal adjustments have ceased to be debatable, they have become imperative In relinquishing the false

values of old habituations Dr Burrow has pointed to new paths and contributed significantly to the understanding of a problem, as yet unsolved, which presses on man from all sides—the problem, namely, of man's own behavior

The Lifwynn Foundation
27 East 37 Street
New York City

(*Wheeler, R. H. The Science of Psychology. New York: Crowell, 1940. Pp. 436.*)

REVIEWED BY DAVID G. RYANS

The title of this book, originally published in 1929 and now revised, has not changed in the course of the author's experience, but the material presented has been increasingly expanded and extended from a systematic viewpoint, and somewhat condensed with respect to argumentative discussion. An attempt is made to provide an elementary text which yields a consistent survey of Gestalt psychology. There is no evidence of an inclination to "write down" the material to an "everyday" level, and the psychological problems and organizing principles are rigorously presented. Early chapters are given largely to general theoretical explanations and to basic laws, while applications to the different areas of psychological investigation are dealt with in later discussions.

Regardless of the personal viewpoint of a reader, he is forced to recognize that this is one of the most thorough-going treatments of general psychology from a distinct systematic standpoint that has been published, and that, in spite of a feeling that there may be gaps between the principles set forth, one is impressed by the experimental results which have been presented to support the "laws" and which may be explained by those laws. The laws, which Wheeler proposes as the groundwork for psychology and which reflect his adherence to a Gestalt explanation, obviously and admittedly lack mathematical precision of statement. They are general and do not, in their present form, enable accurate quantitative prediction, but it is the author's belief that greater precision is certain to follow and that a much more accurate set of predictions will be made available by painstaking experimental research based upon the principles which are here laid down.

Psychology is defined as "that science which studies the problems of human relations with an emphasis upon the interplay of individual human minds or personalities, and their respective structures" (p. 2). This definition is reminiscent of ones given in the first edition of *The Science of Psychology*. There, two definitions were provided, one describing psychology as "a study of conscious behavior and of the conditions under which that behavior takes place" (p. 2), and

the other referring to *human psychology* as "a study, first, of social behavior with its various limited and specialized activities, and second, of those forms of behavior which can be abstracted from it" (p. 2, 3). Behavior, it was explained, was the activity of the organism-as-a-whole. In the course of 11 years Wheeler's definition has become more specific as a description. There is no mention now of behavior as such (and no attempt to define "behavior"). An apparently better chosen term, *personality*, appears in the newer definition, and personality means "the total, organized behavior-pattern of the individual" (p. 69).

For the present and in the immediate future psychology will be most concerned with the development of basic ideas, fundamental concepts, and laws. Later in its history precise demonstrations of principles by means of accurate and detailed experiments may be expected. The general laws applying to psychology apply equally to all scientific branches. In Wheeler's words, "there never has been but one fundamental set of scientific laws of which specific principles in the different sciences are special cases. . . . There is a set of logical constructs or postulates which one can obtain from history, and which, historically, have never been exclusive property of one science but common property of all the sciences" (p. vii). Further, ". . . if psychology is ever to become a predictive science and an ordered body of knowledge at the same time, legitimate principles which transpose from one area in the science to another must be used and these will be the same principles which integrate the human organism with its environment" (p. viii).

It is pointed out early in the discussion that psychology employs both scientific method and scientific laws and may, therefore, be thought of as a science. For Wheeler, scientific method involves both *deduction* (defined as "that process of reasoning, with respect to an observation, whereby a general principle is apprehended to which the observed fact is subordinate") (p. 11), and *induction* ("repeating events or observations under varying conditions") (p. 12).

Three steps are involved in the utilization of scientific method (p. 13-14), namely: (*a*) "the comprehension of laws or principles while making casual observations," (*b*) "the use of these laws in order to plan an experiment, devise apparatus, or to control conditions systematically" (verifying the law by checking it against measurements of the phenomena); and (*c*) on the basis of results thus

obtained, correction and refinement of the laws and restatement, if possible, in precise quantitative language. It is important that the psychologist approach these stages in order. Careful attention must be given to the formulation of basic principles and laws; these must be experimentally verified and tested; and final principles must evolve which are capable of precise statement.

Prediction is thus an end of psychology or of any scientific endeavor. Since the ultimate test of a science is the extent to which its predictions may be considered to be reliable and valid, so the aim of the science of psychology as outlined by Wheeler is "to further the understanding of human nature to the end that *scientific predictions can be made* regarding human behavior" (p. 34). But predictions may be made on different levels, and disfavor should not be conferred upon any science because it is as yet immature and has not advanced to the most precise and quantitative stage. Three stages of prediction are named and it is suggested that these must be arrived at in their respective order under any set of conditions. *Stage One* predictions are those which refer to whether or not an event will occur. This involves "a general, undifferentiated, undetailed comprehension of the whole of a total situation" (p. 35). *Stage Two* predictions indicate the *direction* of an event. These often involve judgments simply of more-than and less-than. *Stage Three* predictions foretell the event and its course in quantitative terms. To illustrate with the example of a falling object, at the first level of prediction we simply predict that the object moves. At the second stage, we are able to predict that the object falls "down" not "up." At the third stage we are able to make predictions as to where a falling object will be at a certain point in space at each instant of time during its fall.

The prediction of events, whether in physics, psychology, or any other branch of science may be made in light of certain general laws. The very general predictions at *Stage One* are made with the aid of knowledge of such laws as the *Law of Field Properties* and the *Law of Determined Activity*. Wheeler states the *Law of Field Properties* as "... wholes exist in their own right over and above the parts or ingredients from which, through closure,[1] they were

[1]Closure—the reorganization of parts or ingredients, resulting in emergence of a new and single whole. Thus, hydrogen and oxygen, two gases, combine to form water, a liquid with properties differing from the original

formed" (p. 26) According to the *Law of Determined Activity,* "the whole regulates the activities of its parts, or, the whole conditions the activities of its parts" (p 26)

The more exact *Stage Two* and *Stage Three* predictions depend upon the uniqueness (and recognition of that uniqueness) of a given event relative to a set of conditions. ("An event is the result of a *whole* set of conditions. A given set of these conditions determines the events uniquely" (p 38) *Stage Two* and *Stage Three* predictions are made possible through knowledge of the operation of the *Law of Least Action* "When action is defined in units of energy multiplied by units of time, movement occurs from one position to another, over the shortest possible path" (p 40) " under a given set of conditions only one course of action will occur, only one process of change will take place, only one path will be followed" (p. 54) (The existence of a goal and the carrying out of goal activities are thus implied by the *Law of Least Action.*)

In all, no less than 18 laws are stated which form the groundwork for the explanations given of various psychological phenomena and processes. In addition to the *Law of Field Properties*, the *Law of Determined Activity,* the *Law of Least Action,* and the *Law of Closure* (already stated), the following principles are necessary to the system here presented:

> *Law of Derived Properties*—"The properties of the parts are derived from the wholes of which they are members" (p 28)
>
> *Law of Individuation*—"Parts come into existence through a division process that can be called individuation" (p 28, 29).
>
> *Law of Action and Reaction*—"A cause could not produce an effect if the effect were a passive victim of the cause The effect must be a cause at the same time the cause is a cause . reaction is the result of 'initiative' both on the side of the cause and on the side of the effect" (p 71).
>
> *Law of Maximum Work*—"For any given set of conditions all the available potential energy of the whole will be expended in the course of maintaining a condition of equilibrium or balance, i e , in preserving its status quo" (p 104). (The Laws

ingredients The *Law of Closure* states, "those forms and shapes of things, or those patterns, will be assumed that under the existing conditions are the easiest to maintain, that is, require the least expenditure of energy per unit of time" (p 56).

of Least Action and Maximum Work are considered supplementary laws.)

Law of Increasing Energy (subsidiary principle under *Law of Maximum Work*)—"The nearer a moving object is to its goal the more kinetic energy it has with which to work" (p. 139) (Similar to Acceleration Law in physics.)

Law of Transposition—"Wholes can be preserved while their parts change" (p 187) (A case of conservation)

Law of Configuration—"A system of energy responds as a whole to a multitude of simultaneous disturbances, and to each disturbance in its relation to others" (p 188)

Law of Reciprocal Change—"With maximum likeness of structure within the whole there goes maximum unlikeness and independence of activity of the parts" (p 233) As parts become differentiated and specialized, making for maximum unlikeness of structure within the whole, increased interdependence of the parts results

Law of Permanence—No special definition is given of this principle, apparently it is assumed to be self-explanatory. By inference it seems to state simply that, other things being equal, an event (response) will tend to be repeated under similar conditions (p 217)

Law Pertaining to Goal—"Whatever object or situation at the time relieves the most tension, or best balances the tensions, (*it*) is the goal which directs the organism's activities" (p 217)

Law of New Insight (a law of perception or observation)—"Within the limits of an organism's mental development, a sudden change in the situation demands new insight and commands its behavior above all else at the time" (p 281)

Law of Observation Range (a law of perception or observation)—"The degree to which an experience is differentiated into discriminations of a refined character is proportional to the size of the field which the observation covers" (p 284)

Law of Field Genesis (a law of perception or observation)—Objects are perceived as evolving wholes so that later perception yields more details and properties, each part being differentiated in relation to the whole (p 298)

Law of Quanta (a law of perception or observation)—"The responding system (organism), in terms of its behavior toward the impinging system (stimuli), structures it, or divides it up, into quanta, which are the sums or increments to which it is responding" (p 320) (This is something of an adaptation

from physics which seems convenient It explains that in observation "wholes" may be obscured, or may not be immediately discerned, while quanta, or parts, of a situation may be perceived due to the discrepancies in the degrees of differentiation between the two systems (organisms and stimuli) impinging on one another)

The "Organismic Laws" stated out of their context may give the casual reader some difficulty as to their understanding and applications. It is pertinent to indicate the manner in which the author has, in *The Science of Psychology,* summarized certain psychological processes in terms of the laws set forth

Learning is described as "a growth and discovery process, expressing itself on the one hand in *insight*[2] into new situations, and on the other hand in the making of new muscular coordinations" (p. 270) Nine laws are applied in describing and summarizing the learning process

(1) Learning involves understanding of the new problem as a whole *Insight* is the *field property* of the response (*Law of Field Properties*)

(2) "The learner acquires control of his responses, in the learning process, under the *Law of Determined Activity*" (p 271)

(3) Details have meaning and purpose derived from their position in the whole (*Law of Derived Properties*), and unless this meaning is discovered learning will not take place

(4) "Goals are essential for learning, and that course toward the goal will be pursued which requires the least expenditure of energy per unit of time" (p 272) (*Law of Least Action,* also, *Law Pertaining to Goal*)

(5) Motivational responses of the organism (increasing the amount of energy available for work), and the resistance of habits to change are examples of the *Law of Maximum Work*

(6) In efficient learning, each detail is learned in its relation to other details (*Law of Configuration*)

(7) "The learning process begins in a general and vague

[2]Insight—"a grasp or understanding of a new situation without (previous) experience in that situation" (p 191) Three criteria of insight are listed (1) *modifiability of behavior*, resulting in correct responses to new situations the first time, (2) *transposability of the general properties of the response* from one problem to another, (3) *configurational character of response,* i e, a response which is made to a total pattern of stimuli

grasp of a problem situation as a whole"—specialization follows as parts become differentiated from the whole (p 273) (*Law of Individuation*)

(8) The unpredictability of early responses and later predictability is an example of *reciprocal change* as structural homogeneity of the whole gives way to heterogeneity of behavior (*Law of Reciprocal Change*)

(9) The nearer the goal of learning is approached, the more energy is available for exertion in reaching or attaining the goal (*Law of Increasing Energy*)

"Personality," according to Wheeler, "is the total organized behavior-pattern of the individual. It is the whole person, as seen behaving from a psychological point of view in all kinds of situations" (p. 69) Organismic laws apply to the topic of personality as follows:

(1) Personality "is a *field property* of the total, organized behavior pattern and is over-summative in character" (p 91) (*Law of Field Properties*)

(2) "The personality is an energy system (*it*) directs its specific responses and adjustments under the *Law of Determined Activity*" (p 92)

(3) "The personality of the individual differentiates from, and within, a variety of overlapping social wholes *Law of Individuation*" (p 92)

(4) "Personality is the result of a growth process in which experiences, habits reaction tendencies *close* into a unitary, single, and complex pattern" (p 92) (*Law of Closure*)

(5) Personalities in action in various situations follow the *Law of Least Action*

(6) Personality develops as the organism and environment interact (*Law of Action and Reaction*)

Not only do organismic laws apply to the learning behavior of the organism and to personality, but they may be employed in an explanation of the nervous system in its relation to behavior as well For example it is maintained (pp 412-416)

(1) The development and functioning of the nervous system (and the entire neuro-muscular system) is in accord with *The Law of Field Properties* (Coghill, Lashley)

(2) Physiological gradients condition the position, direction, and extent of growth in different parts of the nervous system and the organism (*Law of Determined Activity*)

(3) Functions of parts of the nervous system are derived from the position of the parts in the whole (*Law of Derived Properties*).

(4) Development of the embryo and of the nervous system, and development of functions in the cortex, follow the *Law of Individuation*

(5) The *Law of Least Action* supplements the *Law of Individuation* in explaining how specialization of function in the cortex takes place

(6) All or none activity of the nerve impulse is an example of the *Law of Maximum Work*

Similarly, perception makes use of the *Law of Field Properties*, the *Law of Derived Properties*, the *Law of Determined Activity*, the *Law of Individuation*, the *Law of Field Genesis*, the *Law of Least Action*, the *Law of Maximum Work*, and the *Law of Configuration*

The foregoing illustrations provide a representative sampling of the manner in which various psychological topics are systematically presented and described. There would appear to be little value in attempting to point out the shortcomings of the system set forth. Both sides of the questions raised are well known to proponents and opponents alike. A sympathetic reader will find much in which to rejoice as he reviews this highly organized discussion of psychology. One who is less friendly to Gestalt theory will find loopholes and weak points It remains, however, to be said that the treatment has been very painstakingly developed and the basic point of view has been kept in mind at most times. The coverage of material is representative and is generally satisfactory.

Approximately four-fifths as much space is given to textual material in the 1940 edition as in that of 1929 of the *Science of Psychology* The proportionate balance between the major topics in the two edi-

TABLE 1

	Pages of material in per cent 1929	1940
Observational behavior	22 0	21 1
Learning behavior	16 7	16 3
Intelligent behavior	13 3	13 7
Emotional behavior	15 6	13 2
Nervous system in relation to behavior	11 2	13 2

tions is about the same This may be seen readily from the comparisons in Table 1

These topics make up 11 of the 14 chapters of the 1940 edition Ninety-four pages, or approximately 22 per cent of the text, are given to three additional chapters dealing with introductory materials, social behavior, and personality

The style employed by the author is, for the most part, readable and convincing The typology is satisfactory except that there may have been too extravagant a use of italics A list of references and questions for discussion, included at the end of each chapter, provide useful study aids In general, it is the reviewer's belief that the volume marks an advance in the development of Wheeler's viewpoint as originally set forth in the 1929 edition of *The Science of Psychology* The accumulation of psychological data will, after all, put to test the principles and laws laid down by any systematizer It is through such proposals and their acceptance or discard after attempted clarification that progress in any field is made

Cooperative Test Service
15 Amsterdam Avenue
New York City

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i.e., the most recently received books will be found at the end of the list.)

MEGROZ, R L The Dream World New York Dutton, 1939, Pp 319

OGDEN, R M The Psychology of Art New York Scribner, 1938 Pp 291

ANKLES, T M A Study of Jealousy Boston Bruce Humphries, 1939 Pp 111

KARDINER, A The Individual and His Society New York Columbia Univ Press, 1939 Pp 503

WOODWORTH, R S Psychological Issues New York Columbia Univ Press, 1939 Pp. 421

WINKLER, J K Mind Explorers New York Reynal & Hitchcock, 1939 Pp 378

BOODIN, J E The Social Mind New York Macmillan, 1939 Pp 593

GRAY, C T, & VOTAW, D F Statistics Applied to Education and Psychology New York: Ronald Press, 1939 Pp 278

PEAR, T H The Psychology of Conversation London. Nelson, 1939 Pp 171,

TULCHIN, S H Intelligence and Crime Chicago Univ Chicago Press, 1939, Pp 166

JONES, T. D The Development of Certain Motor Skills and Play Activities in Young Children New York Teach Coll, Columbia Univ, 1939 Pp 180

ROSENSTOCK-HUESSY, E Out of Revolution New York Morrow, 1939 Pp 795

LEOPOLD, W F Speech Development of a Bilingual Child. Chicago Northwestern Univ Press, 1939 Pp 188

LAROE, W Parole with Honor Princeton Princeton Univ Press, 1939 Pp. 295

ROETHLISBERGER, F. J, & DICKSON, W J Management and the Worker Cambridge Harvard Univ Press, 1939 Pp 615

JANNEY, J E Laboratory Workbook in Psychology New York Harper, 1939 Pp 154

RATNER, J (Editor) Intelligence in the Modern World John Dewey's Philosophy New York Random House, 1939 Pp 1075

GARRISON, K C The Psychology of Exceptional Children New York Ronald Press, 1940 Pp 351

RASHEVSKY, N Advances and Applications of Mathematical Biology Chicago Univ Chicago Press, 1940 Pp 214

FRAMPTON, M E, & ROWELL, H. G (Editors) Education of the Handicapped Vol 2: Problems New York World Book, 1940 Pp 440

SELLING, L S Men Against Madness New York Greenberg, 1940 Pp 342

ERDELYI, M, & GROSSMAN, F (Editors) Dictionary of Terms and Expressions of Industrial Psychology New York Pitman, 1939 Pp 98

HOHMAN, L. B As the Twig is Bent New York Macmillan, 1940. Pp 291.

KUNKEL, F, & DICKERSON, R E How Character Develops New York Scribner, 1940 Pp 274

GRABBE, P We Call It Human Nature New York Harper, 1939 Pp 120

VALENTINE, W L Student's Guide for Beginning the Study of Psychology (Revised edition) New York Prentice-Hall, 1940 Pp 249

GUILFORD, J P (Editor) Fields of Psychology New York Van Nostrand, 1940 Pp 695

TEAGARDEN, FLORENCE M Child Psychology for Professional Workers New York Prentice-Hall, 1940 Pp 641

DURBIN, E F M, & BOWLBY, J. Personal Aggressiveness and War New York Columbia Univ Press, 1939 Pp 154.

BURLINGAME, L L Heredity and Social Problems. New York McGraw-Hill, 1940 Pp 369.

WHEELER, R H The Science of Psychology New York Crowell, 1940 Pp 436

CANTRIL, H The Invasion from Mars A Study in the Psychology of Panic Princeton Princeton Univ Press, 1940 Pp 228

LATTIMORE, O Inner Asian Frontiers of China New York Amer Geograph Soc, 1940 Pp. 585

GUILFORD, J P, & WALTON, W E. Studies in Elementary Psychology New York: Van Nostrand, 1940 Pp 266

BAKER, H. J The Art of Understanding Boston Christopher, 1940 Pp 400

MERRY, F. K., & MERRY, R V From Infancy to Adolescence New York, 1940 Pp 330.

RHINE, J B, ET AL. Extra-Sensory Perception after Sixty Years. New York. Holt, 1940. Pp 463

WARDEN, C J, JENKINS, T N., & WARNER, L H Comparative Psychology II. Plants and Invertebrates New York Ronald Press, 1940 Pp 1070

WANG, C K A An Annotated Bibliography of Mental Tests and Scales Peiping Catholic Univ Press, 1939 Pp 725

BERNHEIMER, R, CARPENTER, R, KOFFKA, K, & NAHM, M C Art A Bryn Mawr Symposium. Bryn Mawr Bryn Mawr College, 1940 Pp 350

DOOB, L W. The Plans of Men New Haven Yale Univ. Press, 1940 Pp 411

BODE, B. H. How We Learn Boston Heath, 1940 Pp. 308

HULL, C L., HOVLAND, C I., ROSS, R T., HALL, M, PERKINS, D T, & FITCH, F B. Mathematico-Deductive Theory of Rote Learning New Haven Yale Univ Press, 1940 Pp. 329

HARTMAN, G W, & NEWCOMB, T., *Editors* Industrial Conflict: A Psychological Interpretation. New York Cordon, 1940 Pp. 583

KENDIG, I, & RICHMOND, W V. Psychological Studies in Dementia Praecox Ann Arbor, Mich: Edwards, 1940 Pp 211

SIMMONS, R McK A Study of a Group of Children of Exceptionally High Intelligence Quotient in Situations Partaking of the Nature of Suggestion. New York Teach Coll., Columbia Univ., 1940. Pp 112

KATONA, G Organizing and Memorizing New York Columbia Univ Press, 1940. Pp. 318

HUSBAND, R W General Psychology New York Farrar & Rinehart, 1940 Pp 513.

HILGARD, E. R, & MARQUIS, D G Conditioning and Learning New York. Appleton-Century, 1940 Pp 429

GRIFFITH, E. S The Impasse of Democracy New York Harrison-Hilton, 1940. Pp 380

TIFFIN, J, KNIGHT, F B, & JOSEY, C. C The Psychology of Normal People Boston. Heath, 1940. Pp 512.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

December, 1940
Volume 57, Second Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

Entered as second-class matter, August 5, 1897, at the post-office at Worcester, Mass., under the Act of March 3, 1879

Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out.

(1). The subscription price will be $7 per volume, two volumes per year.

(2). The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any zinc or half tone cuts.

(6). The publication charges will be as follows:

4	page signature	$12
8	" "	24
12	" "	36
16	" "	48
20	" "	60
24	" "	72
28	" "	84
32	" "	96
36	" "	108
40	" "	120

These charges are tentative and will be adjusted as justified by the subscription list (especially by the effect of the war in China and in Europe).

(7). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Two volumes annually. Per annum $14.00; per volume $7.00; per half volume unit $3.50. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume | Two volumes per year | December, 1940
Single numbers $4.00 | QUARTERLY | Volume 57, Second Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

DECEMBER, 1940

(Manuscripts are printed in the order of final acceptance)

(over)

Provincetown, Massachusetts

Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass., under the Act of March 3, 1879

Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The Journal of Genetic Psychology, 1940, 57, 249-258

STRANGERS AND FRIENDS AS COMPETITORS AND CO-OPERATORS*

Institute of Child Welfare, University of Minnesota

ALICE JANE PHILP

A. INTRODUCTION

In an attempt to integrate the voluminous literature in the social sciences on competition and cooperation May and Doob (4) pose four summarizing questions: (*a*) Why do individuals compete or cooperate? (*b*) For what things do they compete or cooperate? (*c*) With what persons do they cooperate at all, or compete rather than cooperate, or cooperate rather than compete? (*d*) In what manner do they compete or cooperate?

Our purpose was to compare kindergarten children paired with strangers and with friends at competitive and cooperative tasks. This is one aspect of the third question asked by May and Doob (4).

Very few attempts have been made to study with what persons others will compete or cooperate. However, those previous investigations which have been made suggest that an optimum "degree of likeness" exists between competitors or cooperators. For example Buhler (5) found that babies aged 4-22 months do not compete with each other for a toy if the difference in their age exceeds two and one-half months. According to Von Bracken (6) fraternal twins were less apt to compete than identical twins on tasks in which each knew the other's skill. (The skills of fraternal twins, of course, were less alike than those of identical twins.) Marie Bos (1) studied the "collaboration" (cooperation requiring an exchange of ideas) of pairs of children aged 6-9 and 11-13 on picture matching tasks. She concluded that the intimacy of the children as friends or playmates did not affect the quality of their work although a "too lively personal interest of the children in each other was a stumbling block in the way of fruitful work being done together."

There seem to be three interweaving factors, then, determining

*Indorsed for publication by Florence L. Goodenough, and received in the Editorial Office on August 16, 1939.

with whom people will as a rule compete or coöperate: (*a*) The *likenesses* of the individuals, (*b*) their knowledge of each others' skills, and (*c*) the general conditions under which competition and cooperation occur.

May and Doob (4), in fact, suggest two "levels" on which all competitive and cooperative behavior can be studied. The "social" or "objective" plane may be used when groups of individuals are being investigated. The elements, then, are the "rules, goals, persons and performances." When a few individuals are being studied "relatively thoroughly" their competitive and cooperative behavior may be described in terms of their "discrepancies (between aspiration and achievement), knowledge, attitudes, and skills." The "levels" depend upon each other and upon the time (genetic) element. This study, then, takes place on their "objective level" and varies the "person" element.

B. Subjects

In this investigation 18 kindergarten children and six children of the same age strange to the kindergarten group served as subjects (Table 1). These 24 children were a very select socio-economic sample.

TABLE 1
AGES AND *IQ*'S OF SUBJECTS

	No. boys	No. girls	Median *IQ**	Range of *IQ*	Median age**	Range of age
Kindergarten	8	10	124	100 to 156	5-1	4-6 to 6- 0
"Strangers"	2	4	116	87 to 135	5-1	4-5 to 5-11
Total	10	14	122	87 to 156	5-1	4-5 to 6- 0

*New Revision of the Stanford-Binet administered in 1937-1938.
**Age computed from April 1, 1938, center of testing period.

C. Procedure

The children were paired on five task-situations all of which involved dropping marbles into boxes through small holes. Two tasks were "competitive" (work done on individual boxes to win a paper cut-out), two were "cooperative" (work done on a single enlarged duplicate of the "competitive" tasks to win cut-outs for both members of the pair if the group output was sufficient). One of each type was a steady box on a table (simple), and one of each a box dangling on strings from a scale (more intricate). One other

task required the help of the partner for individual performance (box on a balance pulled down to the child's level by a string).

Subjects were first tested alone twice, without rewards Each child was then tested once with a "stranger," and once with a "preferred" kindergarten child (a child he picked out of the kindergarten group to play games with him). The majority of the subjects, therefore, also served once as a "chosen" child (Table 2).

TABLE 2
DISTRIBUTION OF PREFERENCES

	No of kindergarten children
Not chosen	3
Chosen once	12
Chosen twice	2
Chosen three times	1
Total	18

On the average, then, each kindergarten child participated three times in the main part of the experiment, once with a "stranger," once with a child he chose, and once as a chosen child To equate practice effect of the two groups, "strangers" also participated on the average of three times each Furthermore, each time an unlike sex pairing appeared in the "preference" grouping one was arranged as well among the "stranger" pairings

An attempt was made to equate the number of initial testings for the "stranger" and "preference" pairings This was only approximated, however (Table 3)

TABLE 3
TRIAL ORDER FREQUENCY FOR EACH TYPE OF PAIRING

Trial order	"Stranger"	"Preference"	"Chosen"
1	7	6	5
2	10	5	3
3	1	6	7
4			3
5		1	
Average trial order	1 67	2 17	2 44

At the end of a test period each member of the pair was asked which task (the "competitive" or the "cooperative") he would like to

repeat. Each child's general mode of response during all the tasks was recorded. As the marbles given to each child had been of the same color marble throughout the entire test period, the number of marbles he dropped on the floor could be counted.

D. Results

1. *Comparison of Mean Scores and Variability of "Stranger" and "Preference" Pairings*

From Table 4 it is evident that the mean scores and variability of

TABLE 4

Critical Ratios between Mean Performances with Different Conditions of Pairings

Task	Means	$Diff/\sigma_{diff}$ of means	Sigmas	$Diff/\sigma_{diff}$ of sigmas	r's
Number of marbles put in box during 30 seconds					
Simple "competitive"					
"Stranger" pairings	21.83	.24	4.10	1.38	.35 ± .14
"Preference" pairings	21.50		5.36		
Simple "cooperative"					
"Stranger" pairings	18.00	.69	3.23	.71	.16 ± .15
"Preference" pairings	18.78		3.77		
Time in seconds to put 15 marbles in box					
More intricate "competitive"					
"Stranger" pairings	40.72	.12	7.72	.08	.55 ± .11
"Preference" pairings	40.94		7.62		
More intricate "cooperative"					
"Stranger" pairings	48.00	.36	10.80	.03	.14 ± .15
"Preference" pairings	49.28		10.72		
"Helpfulness"					
"Stranger" pairings	35.06	2.52	7.40	2.50	.41 ± .13
"Preference" pairings	42.16		12.10		

the two types of pairings for the kindergarten children differ very little on any task except "helpful" cooperativeness (Figure 1).

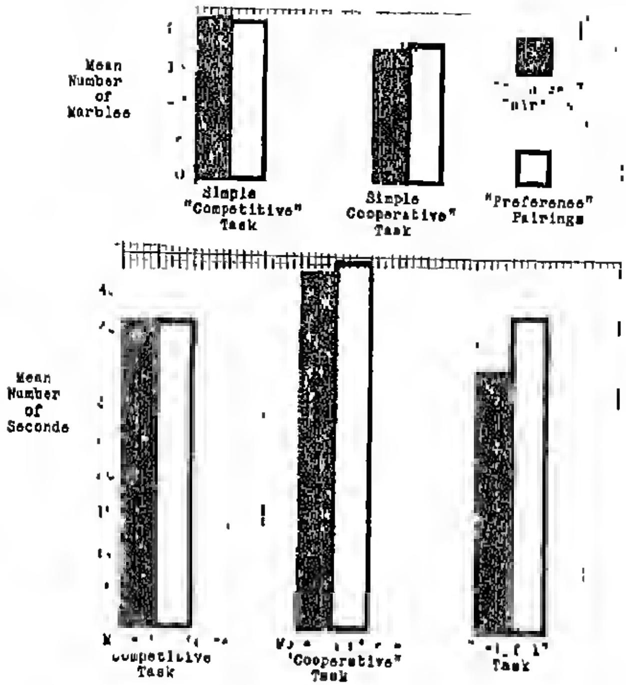

FIGURE 1
COMPARISON OF MEAN PERFORMANCE OF "STRANGER" AND "PREFERENCE" PAIRINGS

2. *Correlations Between Partners' Scores*

Pearsonian *r*'s between partners' scores on each task were computed for the "stranger" pairings and the "preference" pairings (Table 5). The correlations were all low for the "stranger" pairings, but on both of the more intricate tasks (scales) the scores of the "preference" partners correlated .73±.08 and .74±.07.

3. *Correlations of Best "Alone" Scores with the "Stranger" and "Preference" Scores*

When scores on the tasks during the "stranger" and "preference" pairings were correlated with the best "alone" scores on each task, on the "competitive" tasks "stranger" scores correlated more highly

TABLE 5

CORRELATIONS BETWEEN PARTNERS' SCORES*

Tasks	"Stranger" pairings	"Preference" pairings	All partners*
Simple "competitive"	29± 15	05± 16	03± 11
Simple "cooperative"	— 10± 16	22± 16	18± 11
More intricate "competitive"	30± 13	73±.08	.26± 11
More intricate "cooperative"	29± 15	74± 07	55± 08
"Helpfulness"	— 41± 13	20± 15	—.03± 11

*Using Goodenough's (2) formula $r = \frac{\Sigma XY}{N} / \sigma^2$ for obtaining the correlation between interchangeable variables, σ^2 was computed from $2N$ measurements.

TABLE 6

CORRELATIONS OF BEST "ALONE" SCORES WITH THE "STRANGER" AND WITH "PREFERENCE" SCORES ON EACH TASK ($N = 18$)

Task	"Stranger" r's	"Preference"
Simple "competitive"	86± 04	.60± 10
Simple "cooperative"	54± 11	54± 11
More intricate "competitive"	.78± 06	68± 08
More intricate "cooperative"	43± 13	50± 12
"Helpfulness"	38±.14	60± 10

with the "alone" scores than did the "preference" scores (Table 6).

4. *Marbles Dropped During the Experimental Period*

Comparing the two types of pairings again, many more marbles were dropped during the "preference" groupings than during the "stranger" ones (Table 7)

5. *General Response*

There was a very noticeable difference between the types of response noted by the experimenter during the "preference" and "stranger" pairings (Figure 2). "Stranger" pairs were on the whole quiet and "preference" pairs, noisy and talkative

6. *Choice Question*

When the children were asked if they would rather play with the "big" scales ("coöperative") or the "little" scales ("competitive") again, "stranger" pairings chose the "competitive" task more frequently and "preference" pairings the "coöperative" task more often (Figure 3)

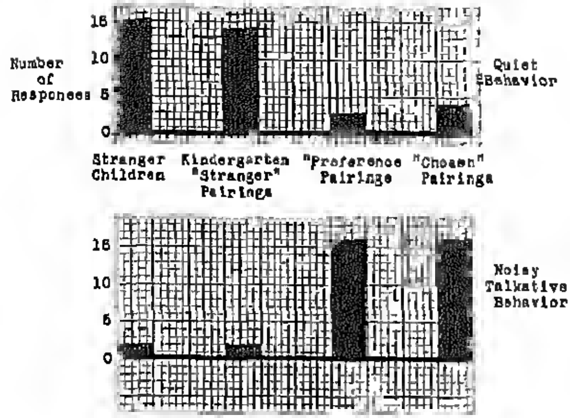

FIGURE 2
General Response

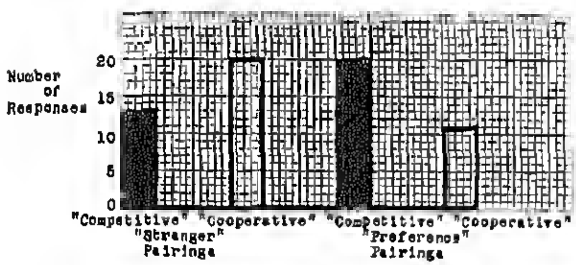

FIGURE 3
Choice Question

TABLE 7
Marbles Dropped During Experimental Period

Pairing	Mean No. marbles	Sigma	r	*Diff.* in means	*Diff.*/σ_{diff}
"Preference"	12.99	15.84	.22±.15	8.43	2.18
"Stranger"	4.56	4.34			

E. Discussion

These results seem to indicate that while no differences exist in mean performance or range of performance between pairings of

"stranger" and "preferred" kindergarten children either at "competitive" or "cooperative" tasks, nevertheless the quality of the response during the two types of pairings is very different Children are more quiet when competing or cooperating with strangers, they drop fewer marbles on the floor, they choose more frequently to repeat the "competitive" task, their scores are less influenced by the performance of their partners, and they perform more as they do when they are working alone What can we infer from these results about a child's attitudes toward children like himself but strange to him as competitors and cooperators?

First, what might be the nature of the stranger-stranger and preference-chosen relationships? (Table 8)

TABLE 8
NATURE OF TWO TYPES OF PAIRINGS

	"Preference"	"Stranger"
1 Knowledge of particular skills	0	0
2 Knowledge of general skills	+	0
3 Likeness—age background, *IQ*	+	+
4 Friendliness	+	0

Second, what possible explanations exist for the similarity in mean performance of the two types of pairings?

1. Perhaps the children did not comprehend the "competitive" or "cooperative" situations, that is, perhaps in both cases the subjects were merely oriented to the task of putting marbles in boxes rather than toward the goal of winning a cut-out Greenberg (3) has shown, however, that before the age of six, 90 per cent of children show competitive behavior This first explanation then, seems unlikely.

2 On the other hand, a "physiological limit" of performance at these simple tasks may have been reached so that increased motivation, had it existed, could not have been measured by mean score. This, too, is unlikely, since we find that preferred partners have influenced each others' scores on several tasks, something that would have been impossible had such a limit been reached

3. The scores were measurements of speed and accuracy combined Perhaps differences would have emerged had either speed or accuracy alone been used as the criterion, or if we had used persistence-time as Wolf (7) did The excitement so noticeable

in the "preference" pairings suggests that speed of work increased, but accuracy decreased in comparison with the "stranger" groupings.

The third question posed by our results concerns the nature of a five-year-old child's response to a stranger. There seem to be three possible types of attitudes, the negative (fearful or hostile), the neutral, and the positive (immediate friendship). We had no evidence of the existence of a "negative" attitude—perhaps this is reserved for individuals very unlike each other in age (size) or background (social status or race). While there were several instances of immediate friendliness among the "stranger" pairings, the majority of pairs seemed to have a neutral attitude toward each other. The strangers treated one another much as if the other person didn't exist—somewhat as a child would treat any inanimate object, behaving as if he were alone. Does this mean that five-year-old children do not yet project sociability to an unknown individual even though that person is like themselves?

For a complete answer to the question we posed at the beginning of this study we need alternately to vary May and Doob's (4) elements on each "level" and vary our four points mentioned in Table 8. Perhaps then we would be in a position to predict a given individual's "social" response when faced with a competitive or cooperative situation involving certain categories of people.

F. Conclusion

When kindergarten children are paired at marble-dropping tasks with strangers and friends as competitors and cooperators, the type of pairing does not influence the mean efficiency or the variability of performance at these tasks, but does affect the quality of the response, stranger pairs being quiet, preference pairings noisy and excited.

References

1. Bos, M. C. Experimental study of productive collaboration. *Acta Psychol.*, 1937, **3**, 315-426.
2. Goodenough, F. L. A short method for computing the correlation between interchangeable variables. *J. Educ. Psychol.*, 1929, **20**, 386.
3. Greenberg, P. J. Competition in children: An experimental study. *Amer. J. Psychol.*, 1932, **44**, 221-248.

4 MAY, M, & DOOB, L Competition and Cooperation New York: Social Science Research Council, 1937

5 *Memorandum on research in competition and cooperation* New York Social Science Research Council, 1937

6. VON BRACKEN, H. Mutual intimacy in twins Types of social structure in pairs of identical and fraternal twins *Charac & Personal*, 1934, **2**, 293-309

7. WOLF, T H The effect of praise and competition on the persisting behavior of kindergarten children Minneapolis Univ, Minnesota, 1938.

199 Arthur Avenue, S E
Minneapolis, Minnesota

The Journal of Genetic Psychology, 1940, 57, 259-277

THE DRAWING ABILITY OF MENTALLY RETARDED CHILDREN*

Department of Psychology, American International College

DOROTHY TILDEN SPOERL[1]

A. INTRODUCTION

This investigation of the drawing ability of retarded children was carried on for three years in a special class of a city school system. Nine children were available all three years, the rest were available for only one of the years with the exception of four children who were available during the first and second years but not the third. During the whole period there were 30 different children in the group.

The range of *IQ* was from 42 to 98, the range of chronological age from seven to fourteen, of mental age from 3-1 to 9-6. Nine were definitely feeble-minded, having an intelligence quotient of 70 or less; 14 were borderline with an intelligence quotient between 71 and 80; three were dull or dull normal, and for four the *IQ* was not available.

There were a total of 418 drawings from the retarded children, the type of drawing to be hereinafter described in the various sections of the study. There were also obtained 56 drawings from normal children who were matched by mental age with six of the retarded children. The drawings were secured during the school years of 1936-1937, 1937-1938, and 1938-1939.

The purpose of the investigation was to study the developmental tendencies in the drawings of retarded children, to compare their performance with that of normal children of like mental age, and to determine if there are specific items which distinguish the work of the retarded child.

The literature on this subject shows a certain amount of overlapping among the characteristics reported as distinguishing the

*Received in the Editorial Office on August 17, 1939.

[1]The author wishes to thank Miss Laura Barker of Miss Barker's School, Springfield, Massachusetts, for furnishing the normal children for this study, and also the school department of the Springfield Public Schools in particular Mr. Frank Doane, and Miss Minnie Esterbrooks.

TABLE 1

ITEMS REPORTED IN LITERATURE AS CHARACTERISTIC OF THE DRAWINGS OF SUBNORMAL AND ABNORMAL CHILDREN

	Subnormal			Abnormal	
	Characteristic	Reported by		Characteristic	Reported by
1	Absence of human figures	Traube			
			1	Head or houses on angle	Traube
2	Arms short in proportion	Rouma			
3	Automatism	Rouma			
			2	Bizarre drawing	Sherman Goodenough
4	*Cliché stereotype*	Rouma	3	*Cliché and fixity of schema*	Earl, Reja
5	Comparable to younger normal	Kerr Goodenough			
6	Slow evolution of stages	Rouma			
7	*Flight of ideas*	Rouma	4	*Flight of ideas*	Earl, Goodenough
8	*Geometric forms*	Traube	5	*Geometric forms*	Earl, Reja
9	Incoherence	Earl			
			6	Individual response	Earl, Goodenough
			7	Inversion (drawing type done by opposite sex)	Earl, Goodenough
			8	Irrelevant emphasis	Earl, Reja
			9	Obscenity	Earl, Reja
10	Regression to inferior stages	Rouma			
11	*Meticulous work*	Rouma	10	*Meticulous accuracy*	Earl
12	*Primitive and mature elements*	Rouma Goodenough	11	*Primitive and mature elements*	Earl, Goodenough
			12	Verbalism	
			13	Brown, blue, black, and violet symbol of depression	Traube, Liss
			14	Long thin houses	Kerr

drawings of subnormal and abnormal children As can be seen in Table 1 five items are reported as characteristics of both sub- and abnormality

Having collected the items that are thought to characterize the drawings of retarded children, the drawings in this study were analyzed to see to what extent these items were prevalent in them The drawings analyzed were those secured in situations where the children were illustrating stories, drawing from free choice, or drawing a definite subject ("what you did yesterday," or "the picnic last week," etc) Four of the items will be discussed separately, the other eight appear in Table 2 with the per cent of drawings showing that item.

TABLE 2

PER CENT OF DRAWINGS SHOWING ITEMS INDICATIVE OF SUBNORMALITY

Item	Feeble minded (75 drawings)	Borderline under ten years of age (85 drawings)	Borderline over ten years of age (65 drawings)	Normal (31 drawings)
Absence of humans	14 7%	5%	12%	45%
Arms short in proportion	26 7%	38%	25%	6 4%
Automatism	24%	14%	8%	3%
Cliché in whole or part of drawing	25.2%	16%	11%	none
Flight of ideas	6%	4%	3%	none
Geometric forms (without meaning)	18%	11%	none	none
Incoherence	28%	12%	8%	22%
Meticulous workmanship	33 3%	37%	65%	9%

From the percentages shown in Table 2 one would tend to conclude that probably the items "absence of humans" and "incoherence" are indicative of immaturity rather than subnormality, since they are found in equal or larger degree among the normal children "Flight of ideas" is found in so small a percentage of drawings that it cannot be said to be substantiated as a characteristic item, yet its higher percentage among the feebleminded than the borderline tends to uphold its inclusion as an item characterizing subnormality "Geometric forms" was found in the feebleminded and one of the

borderline groups, this item however was found in a small percentage of cases, and possibly results from the inclusion of abnormal children in the group. (The writer is studying in another connection the drawings of an abnormal child, which, among other items, show the presence of "geometric forms" to a very marked degree.) The other items: "arms short in proportion," "automatism," "clichés," and "meticulous workmanship" were found in this study to be characteristic of the work of this particular group of retarded children

Four other items remain to be discussed which were not included in Table 2 although they are listed in Table 1 as indicative of subnormality. "Work comparable to that of younger normal children" was found to be true, although the age of the younger children is not consistent This is discussed in more detail later in the paper. Second, "evolution to various stages slow" as reported by Rouma seems to be true since none of the children had progressed to any but the lowest stages of perspective, motion, etc Third, "regression to inferior stages" did not show in the work of these children, tables further on in the paper show a slight but steady progress. Lastly the item "primitive and mature elements combined," was not found in any of the drawings, although it is present to a marked degree in the drawings of the abnormal child mentioned above

It seems to the writer that this is a field in which much further work should be done, establishing further characteristics for the drawings of the subnormal and confirming those characteristics which have been noted.

B. Quantitative Study of the Drawings

The drawings of those children who were available for all three years of the study were scored, with the exception of the drawings which were copies of pictures and the Goodenough test, using the McCarty (6) scales for scoring houses and composition, and the Goodenough test (3) items for scoring the human figures The Goodenough test is not wholly applicable since the situation is not cognate, but it offers a more exact measure than the McCarty scale for human figures This combination of three scores for each drawing seemed more accurate than the use of the Thorndike scale The scores were then averaged and are presented in Tables 3, 4, and 5. Table 3 includes all the children available the first year; Table 4

TABLE 3
AVERAGE SCORES FIRST YEAR OF STUDY

	Average *CA*	Average *MA*	Score on houses	Score on composition	Score on men
Group *A* (24 drawings)	8-5	4-11	9.0 (4)	8.1 (5)	5-5
Group *B* (38 drawings)	8-5	6-4	11.8 (6)	10.4 (6)	6-0
Group *C* (30 drawings)	12-9	9-2	15.7 (9)	16.5 (9)	7-6
Group *D* (18 drawings)	7-5	6-8	13.7 (8)	11.0 (6)	6-2

TABLE 4
AVERAGE SCORES FOR THOSE REMAINING SECOND YEAR OF STUDY

	Year	Average *CA*	Average *MA*	Score on houses	Score on composition	Score on men
Group *A*	I	8-5	4-11	9.0 (4)	8.1 (5)	5-5
(10 drawings)	II	8-11	5-2	11.2 (5)	8.4 (5)	5-9
Group *B*	I	8-5	6-4	11.8 (6)	10.4 (6)	6-0
(19 drawings)	II	8-11	6-8	16.5 (9)	11.4 (6)	6-3
Group *C*	I	11-7	8-10	17.8 (9)	15.9 (9)	6-10
(8 drawings)	II	12-1	9-3	15.6 (9)	11.8 (7)	7-6
Group *D*	I	7-9	7-3	14.9 (9)	11.6 (6)	6-1
(7 drawings)	II	8-3	7-8	14.2 (9)	9.7 (6)	6-11

TABLE 5
AVERAGE SCORES FOR THOSE AVAILABLE ALL THREE YEARS OF STUDY

	Year	Average *CA*	Average *MA*	Score on houses	Score on composition	Score on men
Group *A*	I	7-11	4-10	10.8 (5)	8.9 (6)	5-8
(13 drawings)	II	8-5	5-1	12.5 (6)	8.7 (6)	6-2
	III	9-8	5-10	14.5 (9)	11.8 (7)	7-0
Group *B*.	I	8-2	6-1	12.8 (6)	10.0 (6)	6-0
(24 drawings)	II	8-8	6-5	16.3 (9)	11.9 (7)	6-2
	III	9-11	7-4	16.1 (9)	12.7 (8)	6-11
Group *C* none remain						
Group *D*	I	7-6	6-8	16 (9)	9.1 (6)	6-4
(6 drawings)	II	8-0	7-1	13.9 (9)	9.5 (6)	6-3
	III	9-3	8-3	15.1 (9)	13.1 (9)	7-6

the children available both the first and second year, and Table 5 the children available all three years.

For purposes of comparison the children were divided into four

groups Group *A*, feeble-minded with *IQ* 70 or less; Group *B*, borderline with *IQ* of 71-80 and chronological age under 10, Group *C*, borderline with chronological age over 10, and Group *D* the few dull or dull-normal who were in the class Under the average scores on Houses and on Composition the number in parenthesis is the age level as nearly as it can be determined from the medians given by McCarty (3)

The trends shown in Tables 3, 4, and 5 seem quite definite and surprisingly consistent. For the feeble-minded and borderline-under-10 groups (*A* and *B*) there are consistent advances with each year of the study with the exception of one or two slight regressions in score on houses or composition but giving the same year score The feeble-minded group score definitely in advance of their mental age, the borderline group (under 10) at about their mental age; these results hold for all three types of score on each drawing.

The borderline group (over 10) (*C*) shows a different trend however, first standing still and then regressing slightly in ability as they advance in chronological age Group *D* (dull and dull normal) did about the same type of work as the borderline under 10, whom they surpass but little in mental ability

Since only seven children survived for all three years of the study it seemed wise to average the scores for all the children available in any year (including those considered above as well as those available for only one or two years of the study) This was done so as to see if the general trend would be similar The pictures were scored in the same manner, and the children grouped in the same way The results appear in Table 6.

TABLE 6
ALL RETARDED WITH AVERAGES FOR EACH YEAR

	Year	Average *CA*	Average *MA*	Score on houses	Score on composition	Score on men
Feeble-minded	I	8-5	4-11	9.0 (4)	8.1 (5)	5-5
	II	8-11	5-2	11.2 (6)	8.4 (5)	5-9
	III	8-8	5-3	13.3 (7)	9.9 (5)	6-3
Borderline (under age ten)	I	8-5	6-4	11.8 (6)	10.4 (6)	6-0
	II	8-11	6-8	16.5 (9)	11.0 (6)	6-3
	III	8-4	6-5	14.2 (9)	10.9 (6)	6-7
Borderline (over age ten)	I	12-9	9-2	15.7 (9)	16.5 (9)	7-6
	II	12-0	9-0	16.7 (9)	12.8 (7)	7-6
	III	10-6	7-6	17.0 (9)	14.2 (9)	6-9

The trends as can be seen are very similar as those in the study of the children available all three years The feebleminded somewhat above their mental age level, the borderline (under 10) close to their mental age level, the borderline (over 10) below their level. It is interesting to note too that more than a function of chronological age is at work here, for in Years I and II the feebleminded and borderline (under 10) groups chanced to have exactly the same chronological age, while their performance is quite different.

One other treatment of the material seemed valuable, a study of the average score made at each mental age level for each year. Although the group divided in this fashion gives only a few at each mental age, the consistency of development is of particular interest for that very reason. The curve of development would undoubtedly have been smoother with a more adequate sampling

The results as tabulated in Table 7 show a quite definite progress

TABLE 7

AVERAGES OF ALL RETARDED CHILDREN AT EACH MENTAL AGE LEVEL STUDIED

MA	Year of study	Score on house	Score on composition	Score on men
3-6	III	10.7 (5)	8.3 (4)	4-9
4-6	I	8.3 (4)	6.8 (4)	5-0
4-9	II	7.9 (4)	6.1 (4)	5-1
4-10	III	12.9 (6)	10.4 (6)	5-5
5-3	I	10.9 (5)	8.2 (5)	5-4
5-6	II	13.9 (7)	8.7 (5)	5-5
5-5	III	13.0 (7)	11.3 (6)	6-3
6-3	I	13.2 (7)	11.0 (6)	6-7
6-8	II	16.5 (9)	13.5 (9)	6-8
6-4	III	15.0 (9)	10.4 (6)	7-0
7-4	I	15.2 (9)	12.7 (8)	7-1
7-8*	II	none	13.0 (9)	6-6
7-7*	III	15.9 (9)	13.4 (9)	7-1
8-3**	I	16.9 (9)	14.5 (9)	7-2
8-8**	II	11.2 (6)	13.3 (9)	7-6
8-0**	III	18.2 (9)	8.7 (6)	8-1
9-6**	I	16.4 (9)	17.0 (9)	7-6
9-10**	II	none	11.1 (6)	7-6

*Part of group over 10 *CA*
**All of group over ten *CA*

in drawing score with each increment of mental age. The younger ones do better than their age would lead one to expect (it must be

borne in mind that the very low mental ages however are the result of the inclusion of very low *IQ* ratings), and after the chronological age of 10 is reached there is fluctuation and regression. These figures are highly consistent with those already presented in Tables 3-6

From this quantitative study of the drawings the following trends are observable: (*a*) In retarded children the development of drawing ability is dependent on mental age. (*b*) Feeble-minded children tend to do consistently better work than would be expected from their mental age; the borderline (under age 10) work at about their mental age level, while for the borderline (over age 10) group there is fluctuation and regression This is in line with the findings of Rouma and many others that after the age of 10 drawing ceases to be a form of graphic language and becomes an artistic expression (*c*) Houses reach a maximum score before composition or the drawing of men; composition fluctuates sooner, while the scores on human figures are consistently improved until the chronological age of 10 is reached. Since the houses are for the most part clichés this is in keeping with the tendency to stereotype which does not affect composition or the human figure as strongly.

C. Effect of Type of Drawing Situation

In Years I and III the performance of the children in three types of drawing situations was compared First when the child was merely told "draw what you want," second where the child was told a story and asked to draw an illustration for it, third where a story was told while pictures illustrating it were shown and the children were then asked to draw an illustration for it

In order to reduce the element of chance (that the stories might be in themselves better and more interesting) in the story telling situation the paired stories were either by the same author, or parts of the same book.

The drawings (being in part the same drawings) were scored as in the previous section, and the scores of each child in the three situations compared. If a child had definitely higher scores in any one situation he was judged as having done best in that situation. The results showed that 55 per cent did their best work in the free choice situation, 25 per cent did their best where a story was told while illustrations were shown, and 20 per cent did their best where a story was told without illustrations

However when these scores are broken down into groups the interesting difference of Table 8 appears, again showing a definite

TABLE 8

	Choice	With picture	Without pictures
Feeble-minded	51%	35%	14%
Borderline (*CA* under 10)	62%	35%	5%
Borderline (*CA* over 10)	40%	8%	52%

difference in the ability of those chronologically over the age of 10 They drew best in the story situation when it was not complicated by illustrations.

The trends seem quite clear. (*a*) Feeble-minded and borderline (under *CA* 10) children draw best when they are given their free choice of subject, second best when they are told a story accompanied by pictures (*b*) Borderline (over *CA* 10) children draw best when told a story without illustration, second best in the free choice situation

D Effect of Type of Illustration

During the second year of the study an attempt was made to determine the effect of the type of illustration used in connection with story telling on the subsequent drawings made by the children To determine this four stories were told to the children with accompanying illustrations of four types and the children were then asked to draw an illustration for the story

The following books were used containing the type of illustration described, all had large clear pictures

Realistic Illustration For this a German edition of *Hansel and Gretel* was used (Verlag Jos Scholz, Mainz) which had realistic pictures in soft colors

Decorative Illustration: Helen Sewell's *Cinderella* was used for this (Macmillan, 1934), a book with highly romantic illustrations of a decorative type

Stylistic Illustration The Authors and Artists Guild *Baba Yaga* was used for this, a Pére Castor Book illustrated by Nathalie Parain The trees and people are stylistic, although not so stylized as to lose meaning

Idealized-Realistic For this Ingri and Edgar d'Aulaire's

Children of the Northlights (Viking, 1935) was used, the pictures are realistic but idealized

Two of the stories, *Hansel and Gretel*, and *Cinderella*, were familiar to the children, two, *Baba Yaga* and *Children of the Northlights*, they had not known before

The drawings were scored as in the previous sections of the study, and the average age-score determined for each type of illustration by translating the McCarty scores into ages and averaging the three age-scores for each drawing

Table 9 presents the average scores according to the type of illustration shown in connection with the telling of the stories

TABLE 9
AVERAGE SCORES FOR EACH TYPE OF ILLUSTRATION

	Average *CA*	Average *MA*	*Hansel and Gretel*	*Cinderella*	*Baba Yaga*	*Northlights*
Feebleminded	8-11	5-2	5-8	5-3	7-0	6-2
Borderline (under *CA* 10)	8-11	6-8	7-9	7-2	8-3	7-5
Borderline (over *CA* 10)	12-0	9-0	8-6	8-2	8-5	7-10

It is interesting to note that again the feeble-minded draw definitely above their mental age, the borderline (under 10 *CA*) a little above their mental age; the borderline (over 10 *CA*) definitely below their mental age

Again the feeble-minded and the borderline (under 10 *CA*) do the same type of work, showing the highest scores on their drawings of the unfamiliar *Baba Yaga*, while the borderline (over 10 CA) do their best work with the familiar *Hansel and Gretel*

Bonnie Mellinger (8) found in her study of children's interest in pictures that they preferred the realistic to the conventionalized, while here we find 2/3 of our group drawing better pictures after seeing conventionalized illustrations One possible explanation of this difference (aside from the fact that might be causal that the Mellinger work was a study of normal children) is the work of Freeman (2) who found that 80 per cent of children preferred a saturated to an unsaturated color, while 83 per cent preferred a strong to a weak outline The *Baba Yaga* pictures were strongly

saturated and outlined, while the *Hansel and Gretel* illustrations were weak in saturation and outline It was not possible to find books with the same degree of saturation and outline for the four types of illustration

To see if qualitative judgments would agree with the quantitative judgments 25 judges were asked to judge the pictures of each child from "the dual view of quality and interest." Sixty-six per cent of the judges agreed with the quantitative judgments on the feebleminded, 20 per cent of the judges agreed with the quantitative judgments on the borderline (under *CA* 10), and 100 per cent agreement was found in the case of the borderline (over *CA* 10) In the case of the borderline (under 10) there were a number of children whose pictures were very close in quality

Considering the total effect of the familiar and the unfamiliar stories we find the results of Table 10

TABLE 10

	Average score on the drawings	
	Familiar	Unfamiliar
Feebleminded	5-6	6-7
Borderline (under *CA* 10)	7-6	7-10
Borderline (over *CA* 10)	8-4	8-1

Again we find similar results for the feebleminded and the borderline (under *CA* 10), both doing better drawing with the unfamiliar stories The borderline (over *CA* 10) did slightly better with the familiar stories. It is possible that this may have some connection with the higher tendency to stereotype found with the older children

Each series (i e, all the drawings that were illustrations for one story) were then sorted into groups according to whether they showed definite, vague, or no signs of being influenced by the illustrations seen when the story was being told (this influence was judged by the difference in drawing as compared with the child's usual work) (Table 11).

Apparently as far as definite influence is concerned, not considering the quality of the drawings, there seems to be no great difference except in the case of *Hansel and Gretel* where no draw-

TABLE 11

	Not influenced	Vaguely	Definite influence
Northlights	7	2	4
Baba Yaga	7	2	4
Cinderella	5	3	3
Hansel and Gretel	8	5	none

ings were very definitely influenced It is of interest, however, to enquire as to what type of child was influenced The only child who showed signs of influence in all four of her drawings was one with an *IQ* of 98 It would be interesting to try such an experiment in a group of normal children, and see if their performance was like this of the only normal child in the group Only one feebleminded child showed an instance of influence, and her Goodenough *Drawing Test* scores were always consistently higher than would be expected Of the other children showing influence all but one were chronologically nine or ten (or more) years of age

The influence of type of illustration seen on the drawings produced seems to be: (*a*) The feebleminded and borderline children were more influenced by conventionalized or realistic illustration than by romantic or idealized-realistic (*b*) The borderline (over *CA* 10) children were more influenced by realistic than the other children, conventionalized illustration had the second largest influence (*c*) In general the children who are older chronologically were more influenced than the younger ones

E. The Ability to Copy Pictures in Retarded Children

In the first and the third years of the study the children were asked to copy pictures which were presented to them In the first year three pictures were copied:

> *Cizek.* A drawing from the Cizek school, done by Herta Zuckermann, of a child standing on a hilltop holding flowers while the wind is blowing The original was done by a fourteen-year-old girl, copies were made from a large colored reproduction.
>
> *Cramer:* This was also a large colored reproduction, a scene laid in Brittany with a mother and child outside of a house The human figures compose the major portion of the picture
>
> *Estes:* This was an original pastel done by Harlow Estes, a lake, mountain, pine tree scene in rich blues and greens

In the third year the copies were made from the illustrations of Armer's *The Forest Pool* (Longmans, Green, 1938) The illustrations from this were detached and mounted on black cardboard so that each child had a picture to copy. This procedure was used because the pictures were considerably smaller (8″ by 10″) than those used in the first year study.

The drawings in each series were then rank-ordered according to *their quality as copies* of the original, and this order compared with the rank order of the children on (*a*) their mental age by the school system *IQ*, (*b*) their mental age by the Goodenough *Drawing Test*, (*c*) their chronological age The correlations were determined according to the rank-order correlation formula (Table 12)

TABLE 12

CORRELATIONS

Drawing copied	Correlation with *CA*	Correlation with School *MA*	Correlation with Goodenough *MA*
Cizek	$r = 3645$	$r = 6180$	$r = 6379$
Cramer	$r = 3232$	$r = 3645$	$r = 5176$
Estes	$r = 5075$	$r = 6873$	$r = 5075$
Armer	$r = 3935$	$r = .6379$	$r = 6676$

Except for one instance (copying Estes) the relation between quality of drawing and mental age by Goodenough test is highest; correlation with school *IQ* mental age is second; and with chronological age the lowest. Possibly the purely scenic quality of the Estes picture accounts for this difference.

The Cramer picture was evidently the most difficult for the children to copy, having consistently lower correlations. This is reasonable since both country and costume were unfamiliar

Approximately half of the Cizek and Cramer drawings show a large degree of relation to the original, almost three-fourths of the Estes and Armer drawings Studying the number of children who made copies sufficiently good to be called adequate (i e, including the main details, some sense of the proportions of the original, and some of the "feeling tone") we find that the ease of copying is in this order. Cizek, Armer, Estes, Cramer. If the group of pictures are judged according to which copies show definite influence (are clearly different from the type of drawing ordinarily done) the influence of the originals is in this order, Estes, Cizek, Armer, Cramer

Of the children who made adequate copies 53 per cent were borderline or feebleminded over 10 years of age chronologically, and 6 per cent were dull or dull normal.

We may say then in regard to copying that (*a*) Retarded children in drawing a copy of a picture show a good deal of the original in their copies. Those chronologically over 10 are more successful. (*b*) About half of the children in the groups studied drew pictures which differed recognizably from their usual work. (*c*) The ability to copy is highly correlated with mental age.

F. Comparison of Drawings of Normal and Retarded Children

To test the hypothesis that the drawing of mentally retarded children is in advance of the work that would be expected for their mental age six of the retarded children were matched with six normal children. In each case the mental age of the children matched did not differ more than two months.

Six pictures were secured from each child (both the retarded and the normal); (*a*) "What did you do last Sunday?", (*b*) illustration for story told without illustrations, (*c*) illustration for story told with accompanying illustrations, (*d*) one free choice picture, (*e*) a copy of an illustration from Armer's *"The Forest Pool,"* (*f*) the Goodenough *Drawing of a Man* test.

In order to be sure that the story told was an unfamiliar one, particularly for the normal children who were from a private school,

TABLE 13

COMPARATIVE AVERAGES OF NORMAL AND RETARDED CHILDREN'S DRAWINGS

	CA	*MA*	Average on houses	Average on composition	Average on humans
Normal	4-5	3-10	0	0	0
Retarded	7-1	3-11	12.2	11.5	6-0
Normal	4-0	4-9	7.4	7.1	3-8
Retarded	7-0	4-10	12.9	10.4	5-5
Normal	5-0	5-2	9.0	4.4	4-1
Retarded	10-4	5-4	12.2	12.5	6-1
Normal	6-9	7-4	—	8.4	6-0
Retarded	10-3	7-4	17.6	13.7	7-6
Normal	6-9	7-9	12.0	9.8	5-10
Retarded	10-10	7-8	16.5	14.8	6-1
Normal	6-9	7-10	12.9	11.9	8-6
Retarded	10-1	7-10	13.7	11.6	7-8

the experimenter translated Elsa Beskow's *Resan Till Landet Langesen* (Stockholm: Åhlén and Åkerlund). Half of the story was told without the illustrations, a week later the remainder was told with the illustrations shown to the children.

Table 13 shows the results, giving the average for four drawings (not including the copy of Aimer or the Goodenough test) for each child. With one minor difference (in the case of the last pair) *the retarded child is superior in drawing ability* to the normal child. It is interesting, too, to note that with the normal children there is a steady and consistent advance in drawing ability with each increase in mental age. For the retarded children, especially the two who are over 10 years of age chronologically, there is more fluctuation.

When all of the averages are totaled, and the total average for the normal children is compared with the total average for the retarded the same results are apparent. The average *C.A* of the normal is 5-7, the retarded 9-3. Both have the same average *M.A.*, 6-1. The score on houses for the normal is 8.2 (5), for the retarded 14.2 (9); the score on composition for the normal is 8.3 (5), for the retarded 12.4 (8), the score on men for the normal is 5-7, for the retarded 6-6.

In copying the Aimer drawings the normal children were particularly unsuccessful, complaining more than the retarded ones about the difficulty of the task. Qualitatively judged five of the retarded copies were decidedly better than those of the matched normals, in the sixth case the difference was very slight. Quantitatively judged the difference was more striking. Of the objects in the originals included in the copies the normal children had 12; the retarded 39. The drawing which qualitatively judged was almost like the drawing of the matched child had five elements for the normal child, eight for the retarded one.

In the Goodenough test the normal children all had Goodenough mental ages slightly below the mental age by their school test (Minnesota Preschool, and Terman revision of Stanford-Binet); while with the retarded children the Goodenough mental age was above the mental age except for one child. The average mental age of the normal children was 6-1, their average Goodenough mental age was 6-2, the average mental age of the retarded children was 6-1, their average Goodenough mental age was 7-1. (One normal

child making a very high Goodenough score accounts for the average Goodenough mental age being slightly above their average mental age in spite of most of the scores being slightly below)

Thus in all of the situations the retarded children when compared with normal children whose mental age differed from theirs by not more than two months showed ability which strengthens our theory that: (*a*) The drawing of retarded children is in excess of their mental age (*b*) After the chronological age of 10 the drawing ability of retarded children tends to fluctuate and regress in spite of their mental age. (*c*) The ability to copy is greater in retarded children, perhaps because of their greater chronological age and consequent greater drawing experience

G THE GOODENOUGH TEST IN A CLASS OF RETARDED CHILDREN

During each year of the study the Goodenough test (the drawing of a man) was given to each child in the class The correlations of the *IQ*'s determined by this test and those determined by the City Psychological Laboratory were: Year one. $r = +.561 \pm .115$; Year Two: $r = +.676 \pm .097$, Year Three: $r = +.782 \pm .065$.

Goodenough (3) found her test to correlate with the Stanford-Binet $+.763$, while only one of ours is as high as that. Earl (1), using feebleminded adults, found a correlation of only $+.48$, all of ours being higher than that. The difference seems to be due to our previous findings on the drawing ability of the feeble-minded and borderline children. Of 14 tests on the definitely feeble-minded, 13 received higher mental ages on the Goodenough than by the city laboratory. The fourteenth feeble-minded child made the same *IQ* on both tests Of 13 borderline cases over the chronological age of 10, 9 scored so as to receive Goodenough mental ages lower than their mental age by the city laboratory. Since the groups were different each year in their make-up these facts would tend to explain both the fluctuation and the difference between our correlations and those of Goodenough and Earl.

Put differently: 100 per cent of the feebleminded secure equal or higher mental ages on a drawing test of intelligence, while 69 per cent of the borderline (over 10—chronologically) receive a lower mental age on a drawing test.

It is also interesting to note that if only the Goodenough mental age is considered that five of ten children followed through for two or three years show a consistent growth in mental age although the

IQ fluctuates. Thus they are growing in mental age, but since the growth is slow and uneven the ratio between that and their chronological age is such as to cause *IQ* fluctuation. The other five show growth between the first and second years, or between the second and third years, but have one inconsistent score. Only two over the chronological age of 10 were followed for more than one year, both of these showed a decrease in mental age on the Goodenough test between the first and second years.

McElwee reported an experiment in which profile drawings of subnormal children were matched with profile drawings of normal children whose chronological age equalled the mental age of the subnormals. McElwee concludes (7), "while the profile drawings of the subnormal children showed more mature elements, they also contained more immature elements, such as absence of trunk, attachment of arms and legs at the neck, a row of buttons to represent clothing, and disproportion of parts." None of our matched cases happened to be profile drawings, but in all the profile drawings secured from the retarded children 13 out of 18 failed to score on proportion of arms to body, 17 of the 18 showed consistent disproportion throughout the drawing, and none of the 18 scored more than one or two points on clothing. This seems consistent with McElwee's findings.

Earl (1) matched human figure drawings of feebleminded adults with normal children whose chronological age equalled the mental age of the subnormals. He reports that there is first, a subtle qualitative difference "which can be analyzed into two main factors, firstly a lack of integration whereby undue attention is paid to the details at the expense of the drawing as a whole, and secondly a defective representation of both proportion and position amongst its various elements." He also reports that 55 of his 140 subnormal drawings of the human figure were noticeably asymmetrical. Searching for these three items in our drawings, which were done by retarded children not all of whom were feebleminded, we find that in 10 out of 50 undue attention to detail is shown; in 27 out of 50 there is definite lack of proportion between the parts, and 17 out of 50 show clear asymmetry.

Considering again Rouma's general observation that the arms are shorter in proportion to the body in the drawings of retarded children we find that this appears in 35 out of our 50 Goodenough

test figures (In the general drawings this item was found in 26.7 per cent of the feebleminded, 38 per cent of the borderline-under-10, and 25 per cent of the borderline-over-10, and in only 6 4 per cent of the drawings of the normal children)

Thus we find a high, but fluctuating correlation between the *IQ* of the Goodenough test and that determined by the City Laboratory; this fluctuation probably due to the different performance found in feeble-minded, borderline, and chronologically-over-10-but-feeble-minded-or-borderline children. However the mental age growth between successive Goodenough tests on the same child shows that there is in general consistent growth though the rate of growth is not constant. Significantly large amounts of disproportion, asymmetry, and undue attention to detail were found in the Goodenough test drawings, supporting the findings of Earl and McElwee

H General Conclusions

Our purpose stated at the outset was "to study the developmental tendencies in the drawings of retarded children, to compare their performance with that of normal children of like mental age, and to determine if there are specific items which distinguished the work of the retarded child" The most important developmental tendencies noted were

1 A slow but definite increase in drawing ability with each increase in mental age

2. A marked tendency for the feeble-minded child to draw above the level to be expected from his mental age.

3. A clear tendency toward fluctuation and regression after the retarded child reaches the chronological age of 10 regardless of his mental age.

Comparing the performance of the retarded child with that of normal children of like mental age.

1 The retarded children in general draw better than normal children of like mental age

2 The scores of the retarded children on the Goodenough drawing test are higher than those of the normal children of like mental age

3. The retarded children show greater ability in copying pictures than normal children of like mental age, both when the copies are judged quantitatively and when they are judged qualitatively

No new items characterizing the work of the retarded child

resulted from this study, but our results substantiate previous findings that the work of the retarded child is characterized by.

1 The arms in the drawings of subnormal children are short in proportion to the rest of the figure

2 Clichés are frequent in the drawings of retarded children.

3 The drawings of retarded children show meticulous workmanship

4 Although these children draw in a fashion similar to the work of younger normal children the relationship is not consistent The feeble-minded draw better than those of like mental age, those over 10 tend to draw less well than younger normal children of like mental age

5 Automatism appears in the drawings of retarded children, in particular the drawings of the feeble-minded

6 The drawings of retarded children show lack of proportion between the parts and a large amount of asymmetry.

7 There is some tendency on the part of the retarded child to give undue attention to detail

References

1 Earl, C. J C The human figure drawings of feeble-minded adults *J Psycho-Asthen*, 1932-33, 38, 107-126

2 Freeman, G L, & Freeman, R S The Child and His Picture Book Chicago Northwestern University, 1933

3 Goodenough, Florence L Measurement of Intelligence by Drawings New York World Book, 1926

4. Kerr, M Children's drawings of houses *Brit. J Med Psychol*, 1937, 16, 206-218

5 Liss, E The graphic arts *J Orthopsychiat*, 1938, 8, 95-99

6 McCarty, S Agnes (Ed) Children's Drawings Baltimore Williams & Wilkins, 1924

7 McElwee, E W Profile drawings of normal and subnormal children *J Appl Psychol*, 1934, 18, 599-603

8 Mellinger, B E Children's Interests in Pictures New York Columbia Univ Press, 1932

9 Reja, M L'art malade Dessins de fous *Rev Univer*, 1901, 1, 913-915

10 Rouma, G Le Langage Graphique de l'Enfant Bruxelles Misch & Thron, 1913

11 Sherman, M Interpretation of schizophrenic-like behavior in children *Child Devel*, 1939, 10, 35-42

12 Traube, T La valeur diagnostique des dessins des enfants difficiles *Arch de Psychol*, 1937, 26, 285 309

75 Oak Grove Avenue
Springfield, Mass

The Journal of Genetic Psychology, 1940, 57, 279-288

SPONTANEOUS DELAY FOLLOWING THE APPLICATION OF PUNISHMENT AND REWARD TO STRONG AND WEAK CONNECTIONS*

The Johns Hopkins University and the Baltimore City College

J. M STEPHENS, J A BAER, AND B J. STINNETT

A. INTRODUCTION

The current study is an offshoot of a more comprehensive investigation of the rôle of review or drill in determining the influence of punishment and reward

From a comparison of two previous experiments (3, 4) it appeared that, when a subject is not free to review the correctness of his choice, punishment will not weaken either strong or weak connections. On the other hand, when he is free to make such a review, while still having no influence on weak connections, punishment does weaken strong connections The influence of reward on both strong and weak connections is not significantly affected by restricting the review. These results suggest the possibility that the subjects spend most of their review time in going over strong connections that turned out to be wrong.

The more comprehensive investigation was designed to determine the incidence of the review time—to see if most of the time is spent on strong connections that turned out to be wrong In planning that investigation, however, it became evident that we had to deal with two possible types of delay or review We may have a deliberate review in which the subject is preparing for the re-test due the following day. We may have, on the other hand, a more spontaneous delay due to the element of a surprise rather than to any attempt to prepare for the test on the next day Such a delay should occur even if a subject did not expect a re-test

It happened to be convenient to investigate the latter possibility immediately, while the details of the more inclusive study were being arranged. The current study is concerned only with this spontaneous delay—a delay which occurred when the subjects did not expect a re-test.

*Received in the Editorial Office on August 22, 1939

In order to determine the incidence of the review time, it was necessary to measure how long a subject lingers over each question after he has found out whether his answer is right or wrong

The questions were chosen from the Caldwell and Lundeen superstitions[1] This material was used because it seemed wise to secure a fair proportion of items of the Strong but Wrong type. In ordinary material subjects would tend to be sure of answers that would turn out to be right and this tendency would leave the Strong-Wrong category depleted Consequently we turned to superstitions as erroneous beliefs that might be held with considerable confidence

The questions for each subject were printed on a single sheet of paper. Beside each question there were two places where the chosen answer could be indicated and one place where the degree of certainty could be checked. On the first day the subject worked at an ordinary table, marking his choice and degree of certainty (*guess, hunch,* or *sure*) for each question, but receiving no information as to the correctness of his choice On the second day he worked at the apparatus.

B Apparatus

The apparatus was a modification of that described in a previous publication (4) It consisted essentially of a roller to which the paper was attached and a slot through which one question could be exposed at a time The edge of this slot concealed the marks made on the previous day Each subject operated his own roller and was thus free to linger over a question as long as he chose The paper could not be moved backward so as to reveal questions previously answered

On this second day the subject indicated his choice by punching a stylus through the number of the appropriate answer This completed a circuit through a lamp under the paper and revealed a *W* or an *R*, or merely a blank space, to indicate wrong, right, or no information respectively The subject then went on to the next question as soon as he wished

The wiring diagram of the apparatus is given in Figure 1 The roller (*R*), by which the subject moves the questions past the slot, is held stationary by a brake (*B*) As long as this brake is

[1]We wish to express our gratitude to Dr Otis Caldwell for permission to use this material.

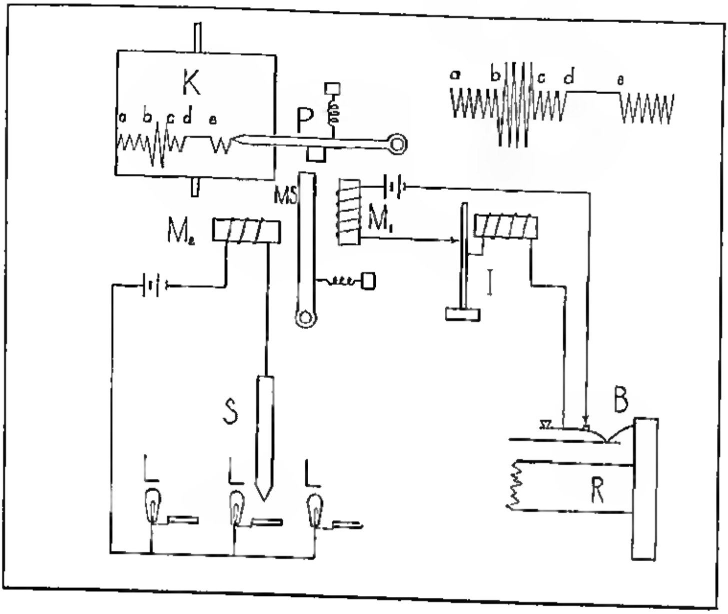

FIGURE 1

applied a current passes through an interrupter (I) and through a signal magnet (M_1) With each vibration of the interrupter the recording pen (P) is pulled down until the projection hits the moveable stop (MS), leaving a record on the continuous paper kymograph (K) This condition obtains while the subject is reading and deciding about the question. As soon as he punches through the paper to indicate his answer, the stylus (S) completes a circuit through the light (L) (flashing R or W) and through a second magnet (M_2) The energizing of M_2 moves the moveable stop (MS) out of the way of the projection on the pen so that the latter can move through a wider excursion By this means the amplitude of the recording pen is increased while its period remains unaffected When the stylus is withdrawn MS falls back into position and the amplitude of the pen is reduced When the subject decides to go on to a new question he has first to release the brake This automatically breaks the circuit through M_1 and leaves the pen tracing a straight line

The time line on the kymograph is redrawn in enlarged form in the upper right hand corner of Figure 1. From *a* to *b* the question is in place for reading, but no choice has been indicated. At *b* the answer is punched and the correctness indicated. At *c* the stylus is withdrawn and at *d* the subject begins to move to a new question At *e* the new question is in position for reading The time used in this study is that intervening between *b* and *d*,—between the time a subject punches and finds out whether he is right or wrong, and the time he goes on to a new question.

This time recording device, an adaptation of one described earlier (2), proved to be very satisfactory

Two possible criteria are available for determining the strength of a choice One may consider as strong all those choices which persist from the first to the second day. One may also use the subject's estimate of certainty. It will be seen that since the estimate is made on the first day it can be applied only to those choices which persist from the first to the second day.

C. Results

Tables 1-4 are based on the records of 30 high school boys and 19 college men and women Table 1 indicates the results obtained

TABLE 1

Average Units of Delay* for All Subjects after Learning that Persisting or Non-Persisting Choices Are Right or Wrong

	Persisting choices		Non-persisting choices	
	Number of choices	Mean units delay	Number of choices	Mean units delay
Choices followed by Right	619	9 1	148	9 5
Choices followed by Wrong	452	10 2	197	10 6

*The unit of delay was an arbitrary measure of approximately 0.45 seconds. Since comparative values only are important, and since conversion into seconds would produce awkward boundaries in the frequency diagrams, the values are left in this arbitrary form

when persistence is used as the criterion of strength or choice

From Table 1 it appears that choices followed by wrong induce more delay than those followed by right There is less difference between persisting and non-persisting choices. The table reveals little, if any, relationship

In Table 2 the data are sorted by the subject's estimate of strength of choice

TABLE 2
AVERAGE UNITS OF DELAY ON THE PART OF ALL SUBJECTS AFTER LEARNING THAT CHOICES OF VARIOUS DEGREES OF CONFIDENCE ARE RIGHT OR WRONG

Subject's degree of certainty	Right		Wrong	
	N	M	N	M
Sure	400	8 9	179	11 2
Hunch	152	9 2	179	9 6
Guess	67	9 9	94	9 3

In Table 2 the Sure-Wrong choices have a slightly longer delay value than that found in any other category, exceeding the next greatest value by 1 3 delay units or 0 6 seconds.

The determination of the significance of differences of this order is a difficult matter The distributions from which these values were obtained are extremely skewed (Figure 2) Any formal estimate of significance would be more dubious than valuable In lieu of such an estimate we give in Figure 2 the frequency polygons of the distributions from which the means in Table 2 were derived. In addition to the frequency polygon for each of the six conditions we have superimposed the polygon of the Sure but Wrong choices upon each of the other polygons after adjusting for difference in area That is to say, we present the Sure but Wrong polygon as it would look if it had the same area as the polygon upon which it is imposed

It will be seen that in every case the values for the Sure but Wrong choices are a little in excess of the values for the other conditions The differences are very slight absolutely and especially slight in view of the wide range of values in all conditions.

A comparison of the distributions leads to the suggestion that the difference lies in the greater tendency of extreme delay values to occur in the Sure but Wrong choices

In another attempt to get at the dependability of this greater delay value for the Sure but Wrong choices, we have subjected the general results to two different analyses The first of these is a comparison of the results from the high school students with those from the college students

In the second breakdown we compare the figures obtained from

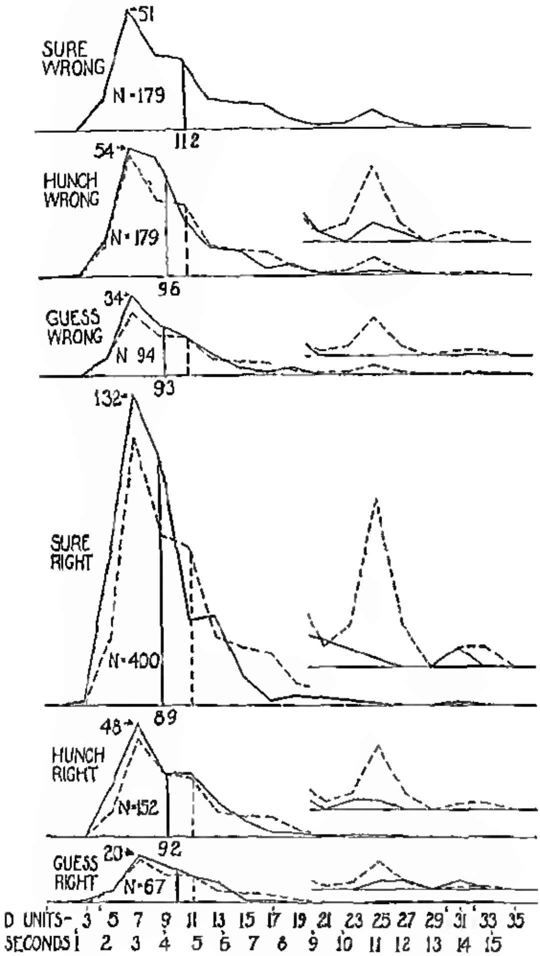

FIGURE 2

the first 20 questions with those obtained from Questions 21-43

From Tables 3 and 4 we see that the general results hold in each of the four sub-tables In every case the Sure but Wrong choices show the greatest delay value.

The next treatment of the data was an attempt to provide still

TABLE 3

	High School				College			
	Right		Wrong		Right		Wrong	
	N	M	N	M	N	M	N	M
Sure	202	8 6	115	10 6	198	9 2	64	12 3
Hunch	85	8 7	113	9 6	67	9 9	66	9 8
Guess	45	9 5	64	9 1	22	10 8	30	9 8

TABLE 4

	Questions 0 - 20				Questions 21 - 43			
	Right		Wrong		Right		Wrong	
	N	M	N	M	N	M	N	M
Sure	161	11 3	107	12 6	239	7 6	72	9.1
Hunch	61	10 6	90	10 8	91	8 4	89	8 6
Guess	33	11 4	54	10 0	34	8 5	40	8 3

another test for consistency and at the same time to eliminate a possible source of error This possible error comes from lumping all choices of all subjects together With lumped data it is possible that the Sure-Wrong choices may be largely from one group of subjects and the Sure-Right choices may be largely from another group It is further possible that the subjects who are very familiar with the material may, on the average, delay less than those who do not know the material so well If, then, we take all those who make "Sure" choices, we will find the well informed, who are also quick, in the Sure-Right category and the poorly informed, who are also slow, in the Sure-Wrong category Hence the Sure-Wrong group will show a greater delay value, not because wrong induces delay, but because people who do know the material well are more rapid than those who do not know it

Although this criticism may not be particularly convincing and although it could perhaps be answered by counter arguments (viz, since on the average the poorly informed should be less confident, we should expect the Guess-Wrong to show the greatest delay), it seemed most pertinent to answer it by re-working the data in such a way that selective grouping could not operate

To apply this correction the results for each subject were treated separately Since each subject made at most only 43 choices, it is obvious that the calculation of the delay value for each of the six categories must be based on a very few cases In fact in the case

of the Guess-Right and Guess-Wrong categories there were many subjects who had no entries at all. In view of this it was decided to combine guess and hunch choices together under the heading of "Not Sure," resulting in the use of only four categories, Not Sure-Right, Not Sure-Wrong, Sure-Right and Sure-Wrong. With this reduction in categories there were only five subjects with one category blank, leaving 44 subjects whose records could be used.

Even with this reduction of categories many of the mean delay values will be based on very few observations (range from 1 to 21), and such means can have little significance in themselves. To overcome this difficulty the means of the four categories were ranked for each subject, as in Table 5.

TABLE 5

	Not sure Right			Not sure Wrong			Sure Right			Sure Wrong		
Subject	*N*	*M*	*R*	*N*	*M*	*R*	*N*	*M*	*R*	*N*	*M*	*R*
Thomas	6	6.5	1	5	12	3	5	10.6	2	4	16	4

This shows that subject Thomas had six choices of the Not Sure-Right type and that the average delay value of these is 6.5 and so forth. We find further than Not Sure-Right has the shortest delay value (Rank 1) and that Sure-Wrong has the longest (Rank 4). Although the individual means may not be very dependable, the ranks based on these means should mean something and the consensus of the ranks should have considerable value. Moreover, this consensus avoids the dangers of lumping data together, since each subject appears with equal frequency (as a rank) in each category.

TABLE 6

The Number of Subjects Whose Delay Value for a Given Category Was First (Shortest), Second, Third or Fourth (Longest)

Category	1st	2nd	3rd	4th	*N*	Av. rank
Not Sure—Right	16	9	15	4	44	2.2
Not Sure—Wrong	13	14	9	8	44	2.3
Sure—Right	11	18	5	10	44	2.3
Sure—Wrong	5	6	14	19	44	3.1
Total*	45	47	43	41		

*These values depart from 44 because ties for first place were both entered under 1st and so forth.

The consensus of these ranks for each condition is presented in the form of a frequency distribution in Table 6.

Table 6 shows that for 19 subjects out of 41 the Sure-Wrong category had the longest period of delay, and for 14 out of 43 subjects it was the next longest

As a further check on the consistency of these results we give in Table 7 similar values for high school and for college subjects separately In this case the numbers have been transmuted into per cents to simplify the comparison of the two groups

TABLE 7

THE PER CENT OF HIGH SCHOOL AND OF COLLEGE SUBJECTS WHOSE DELAY VALUE FOR A GIVEN CATEGORY WAS FIRST (SHORTEST) SECOND, THIRD OR FOURTH

Category	1st	2nd	3rd	4th	N	Av rank
Not Sure—Right						
High Sch	48	15	26	11	27	2 0
College	18	29	47	6	17	2 4
Not Sure—Wrong						
High Sch	22	33	26	19	27	2 4
College	41	29	12	18	17	2 1
Sure—Right						
High Sch	22	44	11	22	27	2 3
College	29	35	12	24	17	2 1
Sure—Wrong						
High Sch.	11	18	30	41	27	3 0
College	12	6	35	47	17	3 2

In Table 7 we find that 41 per cent of high school subjects and 47 per cent of college subjects show the greatest period of delay in the Sure-Wrong condition

The results seem to indicate that when a subject has as much time as he wishes to linger over the correctness of his choice, he will be more inclined to linger over those choices of which he was sure but which turned out to be wrong While there has been no attempt to give the statistical odds in favor of such an hypothesis, its validity is suggested by the consistency with which it is supported in each treatment or analysis of the data Sure but wrong choices show greater delay values than any other type of choice (*a*) when all data are lumped together, (*b*) when the data from high school students and from college students are considered separately, (*c*) when the data from the first 20 choices and the remaining

choices are considered separately, and (*d*) when the trend for each subject is considered separately and those trends averaged

D. Summary

College and high school students (49 in all) were observed with respect to the time they lingered over a question after learning that their choice was right or wrong and before going on to the next question. Some 1400 choices of different types were subjected to four different methods of analysis and in each case the delay values of Sure-Wrong choices (choices of which the subject was sure but which he found out to be wrong) exceeded all other types by a slight margin A comparison of the distributions suggests that the difference may be due to the greater tendency of extreme delay values to appear in the Sure-Wrong category Since the subjects knew there would be no subsequent test, we may regard the delay which appeared as non-deliberative or spontaneous in the sense that it was not a preparation for any specific re-test.

References

1. Caldwell, O W, & Lundeen, G C An experimental study of superstitions and other unfounded beliefs as related to certain units of general science Teach Coll, Columbia Univ, 1932
2. Stephens, J M The conditioned reflex as the explanation of habit formation I The essential factors in the establishment of the conditioned reflex *J Gen Psychol*, 1934, 10, 110-136
3. ——— Further notes on punishment and reward *J Genet Psychol*, 1934, **44**, 464-472
4. Stephens, J M, & Baer, J A The influence of punishment on learning when the opportunity for inner repetition is reduced *J Genet Psychol*, 1937, **51**, 209-217

Department of Psychology
Johns Hopkins University
Baltimore, Maryland

Department of Psychology
Baltimore City College
Baltimore, Maryland

The Journal of Genetic Psychology, 1940, 57, 289-315

CONCEPTUAL RELATIONSHIPS IN CHILDREN THE CONCEPT OF ROUNDNESS*

Normal Child Development Study, Department of the Diseases of Children, Columbia University, and the Babies Hospital

LOUIS LONG

A INTRODUCTION

In the literature of genetic psychology many instances of confusion of concepts, or class names, have been reported It is a common observation that young children apply such a word as "doggie" not only to dogs, but also to cats, to fur neckpieces, and the like The class name at this stage of development seems to be general and is so flexible that it is often practically without meaning to an adult It may, however, be asked whether in fact the species name is at first applied only to many species of the same genus class or whether closer observation would reveal that the same word is used in reference to objects not of the same genus and not possessing any characteristic of the species The latter is certainly a possibility since the inadequacy of the child's vocabulary imposes upon him the necessity of making a few words do service for many objects

It is important in an attempt to understand the genetic development of concept formation, to discover how and when the child delimits the species name and builds up the genus concept Beyond this it is necessary to study the course of development of further hierarchies, with their flexibility and range, and to note with what facility the child can use them

It might be assumed that such problems as the above could best be investigated by studying the verbalizations of children under different conditions Such methods have been employed (3, 9, 11), but it is apparent that this technique will result in the exclusion of children who are too young to verbalize A study of concept formation should not neglect this group if a complete developmental picture is to be obtained Therefore an experimental situation has been devised in which the behavioral indicator is a manual response Data from a study of this type will provide less direct information

*Received in the Editorial Office on August 22, 1939

about certain aspects of the processes involved than an introspective approach might yield; but when subjects are used whose ability to verbalize is inadequate, an analysis of the relation of the behavior to the evoking stimuli permits certain important conclusions to be drawn in regard to the nature of the more subjective components of the response. Some forms of behavior remain unaltered notwithstanding radical changes in the environment In these instances it can be inferred that the stimuli are identical from the point of view of the response mechanism which they set into motion Other variations in the stimulus conditions may call forth a different reaction and these conditions can be said to be non-identical or non-equivalent (4). The determination of the range of equivalence for a particular response will in turn yield information about the nature of the generalization which must be developing as a component of the response

If, in different experimental situations, where a wide range of stimuli is used, the subject is found to respond in a certain prescribed manner to only those factors which are common to all of the situations, then there must be a selective mechanism at work Certain factors of the situation have been abstracted and they have become determiners of the specified response for the subject. Other factors will be ineffective in the production of the response under consideration. The development of the abstraction is a gradual process. During early presentations of the stimuli, the reaction will be first to some and then to other characteristics of the situation Some of these reactions will be reinforced (rewarded) and others will be inhibited (reward withheld) Eventually the significant factor is abstracted from the complex stimulus situation

The learning process by means of which this abstraction takes place is another aspect of the whole problem upon which information should be obtained Learning curves and analyses of the data may indicate a trial and error type of behavior or, perhaps, a solution which can better be classified as "insightful." It may be that a child will consciously formulate and test hypotheses or the response may be made in a completely random and unsystematic fashion Other aspects of the process, more difficult to anticipate, are almost certain to be disclosed

The present article is the report of an investigation of the concept of roundness. The purpose of the investigation was to test

the scope or range of the concept, to discover insofar as possible what children include in the "spherical roundness" concept, and, specifically, to determine whether it includes cylindrical objects and two-dimensional figures. Kluver has pointed out that a practically infinite number of stimulus conditions can be employed in the determination of the "range of equivalent stimuli" (4) Little, however, is known about the conceptual limits or growth of these limits in children This study is an attempt to determine something of the nature of these equivalent stimuli

B. Apparatus

The apparatus consists essentially of two identical and interchangeable compartments in which the objects used as stimuli are placed (Figure 1) The front of each compartment is a one-way

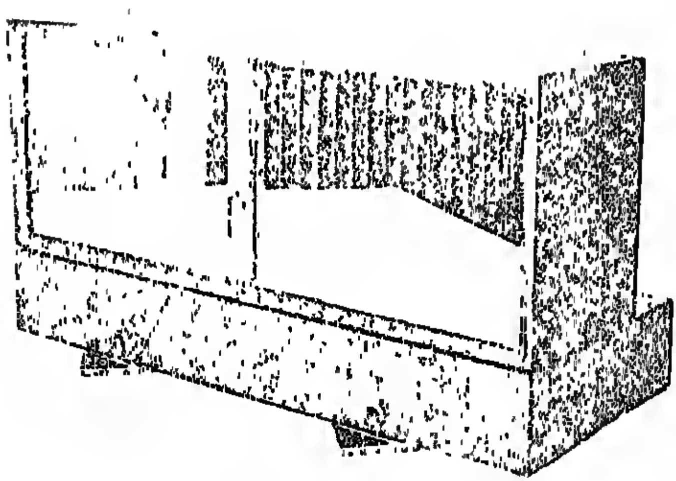

FIGURE 1
Apparatus

mirror screen, which becomes transparent when the illumination is greater inside the box than outside The base into which the compartments plug contains two holes, out of which the reward (a piece

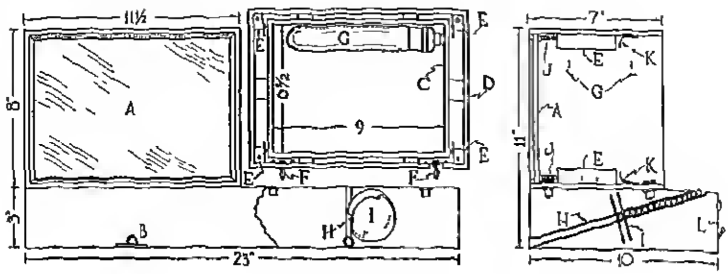

FIGURE 2
Schematic Drawing of Apparatus

A—One Way Vision Mirror
B—Candy Hole and Receptacle
C—Inside Box
D—Outside Box
E—Plungers
F—Connecting Jacks
G—Lamps
H—Candy Tube
I—Feed Discs
J—Plunger Springs
K—Plunger Contacts
L—Control Panel

of candy)[1] rolls into a small receptacle Weak springs placed in the four corners hold the mirrors flush against the front of the compartments. The contact system is arranged so that pressure on the mirror will cause the lights inside the boxes to be turned off and an electric clock to stop At the same time a motor disc system is energized and, if the mirror of the box which contained the positive stimulus was pushed, candy is delivered. If the child responded to the negative stimulus, the motor disc system is not energized, and no candy is received. (By throwing a switch, however, candy may be received from either compartment) Another switch enables the experimenter to select which box is to deliver candy and as this contact is made the lights go on inside the box and the clock starts (An auxiliary switch may be used to start and stop the clock at will.)

Throughout this paper the apparatus described is the one now

[1]The candy used was small, round, and colored, it was a non-sticky variety and dissolved rather readily

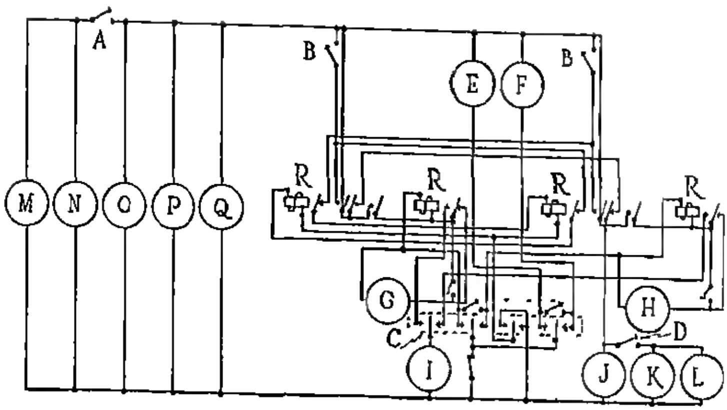

FIGURE 3
WIRING DIAGRAM OF APPARATUS

A—Power Switch
B—Response Switches
C—Selector Switch
D—Timing Switch
E—Left Selection Pilot
F—Right Selection Pilot
G—Left Motor
H—Right Motor
I—Success Pilot
J—Box Lights
K—Timing Pilot
L—Timing Outlet
M—Power Inlet
N—Power Outlet
O—Desk Light
P—Desk Light
Q—Power Pilot
R—Relays

in use From time to time improvements have been made These changes for the most part have had to do with the internal mechanics of the apparatus The one exception to this was the mechanism for the child's response In the early stage of the experiment the response consisted of pressing down a key which extended out from the base of the apparatus The keys, one for each compartment, were directly over the candy holes which were centered for each box It was found that these keys stood out from the rest of the apparatus and consequently the child pressed them spontaneously.

This was a highly desirable feature, but at the same time, the young child seemed to press the keys without looking at the compartment, and so it was decided to attempt to center the child's attention on the compartments by having the response consist of pressing the mirror This system has been in operation long enough to permit the statement that the attempt has been relatively successful, but because the response is not so obvious, one or two demonstrations are sometimes needed When the new system was first installed only one spring was used in each box. This was placed in the upper corner out of sight It is obvious that more pressure is then needed at the bottom of the mirror than at the top This feature has been eliminated by placing springs at each of the four corners of the glass (The springs are concealed within the inner compartment of each box; see Figure 2)

The only parts of the apparatus exposed to the child are the compartments, receptacles, and base containing the candy holes, since curtains completely surround the apparatus and extend from the floor practically to the ceiling

The method seems particularly suited to work with children and has the following advantages

1. Verbal instructions are practically eliminated This should remove the troublesome factor of equating instructions when children of different age levels are used

2 The situation is suitable for children over a wide age range The response is a very simple act, consequently the young child is not penalized by limited motor coordination or lack of verbal ability.

3 The situation is meaningful to the child Everything necessary for the solution of the problem is within his visual range and a close association exists between the response and the reward.

4 Stimuli can be precisely controlled. All types of stimuli can be used, the only limiting factor being the size of the compartments. The stimuli are removed from the child's vision as he responds. All adjustments and arrangements of the stimuli are made outside the visual range of the child.

5 All secondary cues are eliminated *E* is behind the curtains All auditory, visual, tactual and olfactory cues are ruled out. The apparatus is perfectly silent during the period of exposure of the objects Only after the child responds does the apparatus make a noise as candy is being dispensed There is also a click of the

relay as the light is turned on for the next trial, but this has been found very convenient because the child soon learns to use it as an indication that the machine is ready to work again. There are no visual cues since the compartments are identical and interchangeable. The same amount of pressure is required in responding to a compartment regardless of whether or not it contains the positive stimulus, so that no tactual cue operates as a differential. Since candy is always kept ready to be dispensed from both compartments no olfactory cue is present

6 The stimuli are removed from the child's vision as soon as he responds.

7. No screen is superimposed in front of the child between trials and thereby any extraneous disturbing factor is avoided A one-way mirror prevents the child from observing the experimenter, and at the same time, affords some children entertainment during the recording and arranging of objects

8 *E* can observe the child, without the latter being aware of it, by a one-way system of mirrors

9. Motivation is extremely good.

C PROCEDURE

The child is brought into the room and is seated before the apparatus *E* goes behind the curtains (answering any questions *O* asks about where *E* is or what he is doing in a vague manner, e g, *"I am working"*) The lights are then turned on and a ball and a block (Condition 1) are exposed, one object in each compartment The child's reaction is noted through the one-way vision system Usually a great deal of exploration, including questions or statements about the objects, occurs as soon as the lights go on Sometimes the mirrors will be touched in this process of initial investigation *E* attempts to turn questions such as: what shall I do, how does it work, etc, back to the child by asking, for example, how do *you* think it works? If after a period of about five minutes the child has still not pressed the mirror, *E* presses the mirror of the compartment containing the positive stimulus[2] *O* is usually greatly surprised by the lights going off and the candy coming out The customary reaction is *"Do it again"* In some cases it has been

[2]In scoring, these trials are considered as incorrect responses

TABLE 1
STIMULUS PAIRS AND CRITERIA

Condition No		Criterion
	Type S stimuli	
1	Red rubber picture ball (radius—3.4 cm) vs Natural wooden rectangle (5.4 x 5.4 x 2.7 cm)	20 consecutive Rs to Pos S
2	Red rubber picture ball (radius—3.4 cm) vs Cream-colored wooden triangle (14.5 x 6.8 x 2.2 cm)	10 consecutive Rs to Pos S
3	White rubber golf ball (radius—2.1 cm) vs White creased paper drinking cup (6.5 cm high, R of top—2.8 cm., R of bottom—2.0 cm)	10 consecutive Rs to Pos S
4	White and red striped marble (radius—0.8 cm) vs Natural wooden rectangle (5.3 x 2.7 x 2.7 cm)	10 consecutive Rs to Pos S
5	Orange vs Drinking glass (9.5 cm high, R of top—3.6 cm, R of bottom—3.1 cm)	10 consecutive Rs to Pos S
6	Small red wooden ball (R—2.25 cm., V—48 cu cm) vs Small red wooden cube (L—3.7 cm, V—51 cu cm.)	10 trials
7	Small green wooden ball (R—2.25 cm, V—48 cu. cm) vs Large green wooden cube (L—7.2 cm; V—373 cu cm)	10 trials
8	Large green wooden ball (R—4.4 cm; V—357 cu cm) vs Small green wooden cube (L—3.7 cm, V—51 cu cm)	10 trials
9	Large red wooden ball (R—4.4 cm, V—357 cu. cm) vs Large green wooden cube (L—7.2 cm, V—373 cu. cm)	10 trials
10	Large red wooden ball vs Large red wooden cube	10 trials
11	Large red wooden ball vs Small red wooden cube	10 trials
12	Large red wooden ball vs. Small green wooden cube	1 trial
13	Large green wooden ball vs Large red wooden cube	1 trial
14	Large green wooden ball v Small red wooden cube	1 trial
15	Small red wooden ball vs. Large red wooden cube	1 trial
16.	Small red wooden ball vs Large green wooden cube	1 trial
17	Small red wooden ball vs. Small green wooden cube	1 trial
18	Small green wooden ball vs Large red wooden cube	1 trial
19	Large green wooden ball vs Large green wooden cube	1 trial

TABLE 1 (*continued*)

Condition No		Criterion
20	Small green wooden ball vs Small red wooden cube	1 trial
21	Small green wooden ball vs. Small green wooden cube	1 trial
21A	Background varied Y vs N (No background), N vs R; B vs Y; G vs N, R vs G (Each pair presented once)	5 trials
	Type C stimuli	
22	Natural wooden disc (R—3.7 cm depth—2.0 cm) vs White rectangular paper box (4.9 x 4.9 x 2.3 cm)	10 trials
23.	Round brass metal box (R—2.0 cm, depth—1.6 cm) vs 8-sided green glass pill bottle (5.5 x 3.4 x 2.3 cm)	10 trials
24	Red rubber circular sponge (R—3.5 cm; depth—3.7 cm) vs White rubber rectangular sponge (5.7 x 5.4 x 4.6 cm.)	10 trials
25	Black perforated typewritten ribbon holder (R 2.7 cm, depth—1.6 cm) vs Rectangular wooden match box (5.9 x 3.7 x 1.8 cm)	10 trials
26	Natural wooden circular peg (R—1.8 cm, length 5.5 cm) vs Yellow wooden (six-sided) rectangular pencil (sides—0.3 cm, length 15.0 cm with point incl)	10 trials
27	Aluminum circular box (R—1.6 cm., length—4.2 cm) vs. U-shaped black cover (4.4 x 2.9 x 2.0 cm)	10 trials
28	Yellow circular roll of sealing paper (R—4.0 cm, R of hollow portion—1.8 cm, depth—5.1 cm) vs Rectangular glass lid (top—5.1 x 5.1 cm, bottom—4.9 x 4.9 cm, depth—1.6 cm)	10 trials
29	Circular glass bottle (R—0.9 cm; length—8.2 cm) vs Rectangular glass bottle (5.3 x 2.2 x 2.2 cm)	10 trials
30.	Circular paper outline (R—4.1, R of hollow portion—3.8; width—2.5 cm) vs Rectangular paper outline (6.5 x 6.5 x 2.5 cm; hollow portion—6.3 x 6.3 cm)	10 trials
	Type T stimuli	
31	Two-dimensional figures (4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15-sided figures) and a circle, area of all—48 sq cm (approx) R—2.25 cm Each figure paired with every other figure once in haphazard order	78 trials
32	Two-dimensional figures (same as Cond 31) Each many-sided figure paired with circle five times	60 trials

found necessary to give another demonstration on the second trial.

The amount of assistance varies from child to child. The complete elimination of this variation would be highly desirable, but with children differing markedly in such characteristics as exploratory urge, initiative, and persistence, a standard procedure in regard to assistance would not solve the problem. The following procedure was employed to eliminate the effects of the variation. An arbitrary criterion of consecutive responses (20 in this instance) to the positive stimulus has been set up and the data on the trials necessary to reach this level of performance are not considered in the present analysis so that the variation in instruction will not effect the final result. Thus, the first situation (a ball and a block) serves as an orientation period and in a sense the experiment begins only after 20 correct responses have been made.

At this point the pair of stimuli listed as Condition 2 (see Table 1) are exposed. In this condition the criterion is 10 consecutive correct responses. When this criterion has been reached Condition 3 is presented to the child; the criterion is again 10 consecutive responses to the positive stimulus. Table 1 presents a list of the conditions along with a description of the objects. The criterion for each condition is also listed in this table. It will be noted that for Condition 6 and the conditions following it, *"10 trials"* or *"1 trial,"* is listed. For these conditions each pair of stimuli was presented exactly the number of times designated, regardless of whether the child responded to the positive or to the negative stimulus. The positive or rewarded stimulus is always listed as the first member of the pair. The negative stimulus was not rewarded. (An exception to this will be discussed later.)

Another factor in the procedure that should be mentioned is the conversation between the child and the experimenter. The latter encouraged this conversation and his part in it consisted mainly of asking an occasional question. Particular care was taken to avoid any standard questions or remarks that might serve as cues to the child.

The position of the positive stimulus was shifted from left to right according to Gellermann's prearranged chance order (2).

An attempt was made to keep the intervals between experimental sessions constant, but in several instances illness made this impossible. The children were tested on Monday, Wednesday, and Friday.

The time of the day was constant for any one child, but varied from child to child. It was found that the relationship between the time of experimenting and the last meal was unimportant for all children except the very youngest These children were tested before their mid-day meal.

As a result of preliminary work it was decided to make no attempt to keep the number of trials per session constant The only constant factor was the maximum number of trials per day, which was set at 40 Ordinarily the testing period took about 30 minutes It was found that on some days the child was perfectly content to run through 30 or 40 trials while on other days he might become restless and inattentive after 15 or 20 trials. When the latter condition existed the experimenting was terminated for the day.

D. Stimuli

The stimuli are listed in Table 1 The discussion will be facilitated by breaking up the list into three parts The first group (Conditions 1 through 21*A*) consists of three-dimensional, spherical objects (Type *S*) These stimuli can be further divided into two groups In the first sub-group (Conditions 1 through 5) the positive stimuli are all perfect spheres, except the fifth, which is not so symmetrical as the other four Two of the negative stimuli (cup and glass) might well be considered round also The second sub-group (Conditions 6 through 21*A*) consists of eight objects which represent all possible combinations of the three variables used, viz., size (small, large), color (red, green), and shape (ball, cube) These objects are arranged in 16 pairs and within four of these pairs (Conditions 6, 10, 19, 21), shape is the only variable which differs. In Condition 6, for example, the objects are alike in size, color, brightness, and material, and differ only in shape.

These objects were the last members of the first group and were introduced so that the concept of spherical roundness might be clearly formulated before a shift was made to cylindrical and two-dimensional objects. This technique does not prove that it is roundness alone to which the child is reacting, for, as Kluver (4) has pointed out, the response is dependent upon a constellation of stimuli and it is possible that had some of the identical characteristics of the stimulus (e g., brightness, material) been varied a breakdown

in the discrimination might have occurred in spite of the continued presence of the roundness variable. The procedure has, however, eliminated the operation of two variables (size and color) which might be expected to be especially effective in determining the responses of children of the given age level.

In Condition 21*A* the factor of background is varied in an exploratory fashion. The pair of objects used in Condition 21 is again presented five times (see Table 1), but instead of the objects resting on the floor (painted white) of the box, a piece of colored paper, covering the entire bottom of the box, is inserted between the floor and the object. The colors presented were red, yellow, blue, and green.

The second group of stimuli (Conditions 22 through 30) comprises familiar objects, the positive stimuli (Type *C*) being cylindrical, but not spherical. Several negative stimuli (Conditions 23 and 29) which from certain aspects might also be considered round, were introduced.

The third group of stimuli (Conditions 31 and 32) is composed of figures cut from black cardboard (Type *T*).[8] They are alike in area, color, and material. They were presented vertically. The group consists of a circle and 12 many-sided figures, the number of sides ranging from four through 15. In Condition 31, these 13 figures are presented so that each figure is paired once with every other figure, making in all, 78 combinations presented in a standard haphazard manner. In Condition 32, one member of the pair is always the circle, the other stimulus being one of the many-sided figures. The procedure was to present the circle paired with each of the other figures five times, starting with the four-sided one and increasing the number of sides by one after each set of five trials.

E. Subjects

Thirteen subjects took part in this experiment. The age range was from three years to six years, three months. (See Table 2 for age, sex, *IQ* and length of experimental period for each child.) The children are members of the Normal Child Development Study and are all normal and healthy. The socio-economic status of their

[8]These figures may be considered two-dimensional in spite of the fact that they were presented against a background. This assumption was checked by using figures painted on cardboard.

TABLE 2
AGE, SEX, INTELLIGENCE QUOTIENT, AND LENGTH OF EXPERIMENTAL PERIOD OF SUBJECTS

Subjects	Sex	Interval of expt, mo da	Age at mid pt of expt, yrs mo	IQ
E D	F	7-29	3-0	100
RoM	M	2-18	3-5	118
A R	M	0-24	3-6	118
L D	M	4-1	4-10	105
J C	M	1-13	4-11	102
T P	M	4-3	4-11	121
B B	F	8-11	5-4	137
C C	F	1-9	5-4	134
F M	F	3-15	5-5	98
M M	F	4-17	5-5	104
M W	F	0-16	5-10	96
M H	F	2-18	6-0	115
A D	M	3-23	6-3	123

parents is about that of the average New York family These children had been subjects in experiments conducted by other members of our staff as well as by the writer They are, therefore, accustomed both to experimental routine and to the present experimenter.

F Results

The percentage of responses to the positive stimulus in each condition is presented in Table 3 for the 13 children who learned the first situation Each child's score, along with the average for the entire group, is given This table suggests a relation between age and percentage of correct responses. A rank difference correlation of 64± 12 was found between chronological age and the mean per cent for all conditions. Similarly a correlation of 58± 13 was found between this mean per cent and mental age The table shows that in 11 instances out of 416 the percentage of responses to the positive stimulus is less than 50. In 26 instances (including the 11 cases above) the percentage is less than 70, and in 64 instances it is less than 90. It follows that the positive stimulus was chosen 50 per cent of the time, or more, in each of 405 out of a possible 416 situations, and 90 per cent of the time, or more, in each of 352 situations. On the basis of this last proportion, the general statement may then be made that in all but 15 per cent of the situa-

TABLE 3
PERCENTAGE OF RESPONSES TO POSITIVE STIMULUS

Conds	ED	RoM	AR	LD	JC	TP	CC	BB	FM	MM	MW	MH	AD	Av of all sub	No of trials for ea Cond
2	100	91	100	100	100	92	100	100	100	100	100	100	100	99	
3	86	81	38*	100	95	94	100	81	92	91	91	100	100	88	
4	100	92	100	100	100	100	100	100	100	100	100	100	100	99	
5	100	92	100	100	100	100	91	100	93	93	100	100	95	97	
6	90	100	80	100	100	100	100	100	100	100	100	100	100	98	10
7	60	100	80	70	100	100	100	100	100	100	100	100	100	93	10
8	100	100	100	100	100	100	80	100	90	90	100	90	90	95	10
9	100	60	90	100	100	100	100	100	100	100	100	100	100	96	10
10	70	100	90	100	100	100	90	100	100	100	100	100	100	96	10
11	90	100	100	100	100	100	100	100	100	80	100	100	100	98	10
12	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1
13	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1
14	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1
15	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1
16	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1
17	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1
18	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1
19	100	100	100	100	100	100	100	100	100	100	100	0	100	92	1
20	0	100	100	100	100	100	100	100	100	100	100	100	100	92	1
21	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1
21A	100	100	100	100	100	100	80	100	80	100	100	100	100	97	5
22	20	40	80	100	80	80	80	100	90	100	100	70	80	78	10
23	50	100	100	90	80	90	70	90	90	100	100	80	100	88	10
24	10	100	100	100	90	60	90	100	90	100	100	70	100	85	10
25	10	60	90	100	90	90	80	100	90	100	100	90	100	85	10

TABLE 3 (*continued*)

Conds	ED	RoM	AR	LD	JC	TP	CC	BB	FM	MM	MW	MH	AD	Av of all sub	No of trials for ea Cond
26	30	40	80	100	100	100	100	80	90	100	100	90	100	85	10
27	90	60	100	100	100	100	90	100	100	100	100	100	100	95	10
28	70	70	90	100	100	100	100	90	100	100	90	80	100	92	10
29	60	40	90	90	80	100	100	100	100	90	90	90	100	87	10
30	70	60	90	100	100	90	90	100	100	100	100	60	100	89	10
31	50	55	69	69	76	87	73	86	83	86	86	68	78	74	78
32	50	45	93	100	95	100	95	100	100	100	100	88	100	90	60
														Total	Per cent of all Conds for all sub
No of Percentages Below 50	5	4	1	0	0	0	0	0	0	0	0	1	0	11	3
No of Percentages Below 70	10	9	2	1	0	1	0	0	0	0	0	3	0	26	6
No of Percentages Below 90	14	11	6	2	4	3	6	3	2	2	1	8	2	64	15

*Criterion of 10 consecutive responses to positive stimulus not reached in 40 trials
$N=416$

TABLE 4
NUMBER OF RESPONSES TO NEGATIVE STIMULUS ON FIRST TRIAL OF EACH CONDITION

	SUBJECTS														
Conds	ED	RoM	AR	LD	JC	TP	CC	BB	FM	MM	MW	MH	AD	Total	Group Total
							Type S stimuli								
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
3	1	1	0	0	0	0	0	0	0	1	1	0	0	4	
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
5	0	0	0	0	0	0	1	0	0	0	0	0	0	1	
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
8	0	0	0	0	0	0	1	0	0	1	0	1	0	3	
9	0	1	1	0	0	0	0	0	0	0	0	0	0	2	
10	0	0	0	0	0	0	1	0	0	0	0	0	0	1	
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
19	0	0	0	0	0	0	0	0	0	0	0	1	0	1	
20	1	0	0	0	0	0	0	0	0	0	0	0	0	1	
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
21A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
							Type C stimuli								
22	1	0	0	0	1	0	1	0	1	0	0	1	0	5	
23	1	0	0	0	1	1	1	1	1	0	0	1	0	7	
24	0	0	0	0	1	0	0	0	1	0	0	0	0	2	
25	1	1	0	0	1	0	1	0	1	0	0	1	0	6	
26	0	1	1	0	0	0	0	1	0	0	0	1	0	4	
27	0	1	0	0	0	0	0	0	0	0	0	0	0	1	
28	1	0	1	0	0	0	0	0	0	0	1	1	0	4	
29	1	1	1	0	1	0	0	0	0	0	1	1	0	6	
30	0	0	0	0	0	0	1	0	0	0	0	1	0	2	
							Type T stimuli								
31	0	0	0	0	0	0	0	0	0	0	1	0	0	1	
Total	7	6	4	0	5	1	7	2	4	2	4	9	0		51

$N = 403$

tions the positive stimulus continued to be selected by those subjects who had mastered Condition 1 This consistency of performance implies a realization of the concept involved in the experiment

Additional evidence which seems to demonstrate the satisfactory establishment of the concept can be gained from an analysis of the first responses made after a change of stimulus objects Table 4 presents the number of responses to the negative stimulus on the first trial of each condition This table shows that the negative stimulus was responded to in only 51 instances (13 per cent) on the first trial The significance of the first response to a new situation as an indicator of the effect of previous training has been stressed by Warden and Rowley (10) in connection with transposition experiments Although this is not a typical transposition experiment, the importance of the first reaction is not thereby diminished This response is significant because it is the only reaction to a specific stimulus situation which is free from the effect of a previous response to that situation That is, the second, and all succeeding, reactions will be affected by the correctness or incorrectness of the first Since the round, or roundest, object was selected the first time in 352 instances out of 403, we have some indication that the concept of roundness determined the subject's reactions

It will be recalled that eight equated objects were presented in Conditions 6 through 21*I* Due to the objective similarity of this group it is presented as a unit in Table 5 Such treatment was planned beforehand and consequently, as noted above, some conditions were presented only once By a grouping of the data, the chance factor which would have made the results from those conditions presented only once of dubious value, has been counteracted Table 5 is divided into three parts corresponding to the three types of stimulus (perfectly spherical objects, cylindrical objects, and two-dimensional figures) (See Figure 4 for a graphic presentation of Table 5) This table shows clear-cut differences in the results for the three groups. The average group percentage of responses to the positive stimulus for all children is highest for Type *S* (spherical objects) lowest for Condition 31 of Type *T*[1] (two dimensional objects) and midway for Type *C* (cylindrical objects) and Condi-

[1]In this discussion Condition 31 will be considered the main condition of Type *T* stimuli The apparent discrepancy between the two groups will be discussed below

TABLE 5

PERCENTAGE OF RESPONSES TO THE POSITIVE STIMULUS

Conds	ED	RoM	AR	LD	JC	TP	CC	BB	FM	MM	MW	MH	AD	Av of all subjects
							SUBJECTS							
							Type S stimuli							
2	100	91	100	100	100	92	100	100	100	100	100	100	100	99
3	86	81	38*	100	95	94	100	81	92	91	91	100	100	88
4	100	92	100	100	100	100	100	100	100	100	100	100	100	99
5	100	92	100	100	100	100	91	100	93	93	100	100	95	97
6-21A	87	95	92	96	100	100	95	100	97	96	100	97	99	96
Average	95	90	86	99	99	97	97	96	96	96	98	99	99	96
							Type C stimuli							
22	20	40	80	100	80	80	80	100	90	100	100	70	80	78
23	50	100	100	90	80	90	70	90	90	100	100	80	100	88
24	10	100	100	100	90	60	90	100	90	100	100	70	100	85
25	10	60	90	100	90	90	80	100	90	100	100	90	100	85
26	30	40	80	100	100	100	100	80	90	100	100	90	100	85
27	90	60	100	100	100	100	90	100	100	100	100	100	100	95
28	70	70	90	100	100	100	100	90	100	100	90	80	100	92
29	60	40	90	90	80	100	100	100	100	90	90	90	100	87
30	70	60	90	100	100	90	90	100	100	100	100	60	100	89
Average	46	63	91	98	91	90	89	96	94	99	98	81	98	87
							Type T stimuli							
31	50	55	69	69	76	87	73	86	83	86	86	68	78	74
32	50	45	93	100	95	100	95	100	100	100	100	88	100	90

*Criterion of 10 consecutive responses to positive stimulus not reached in 40 trials

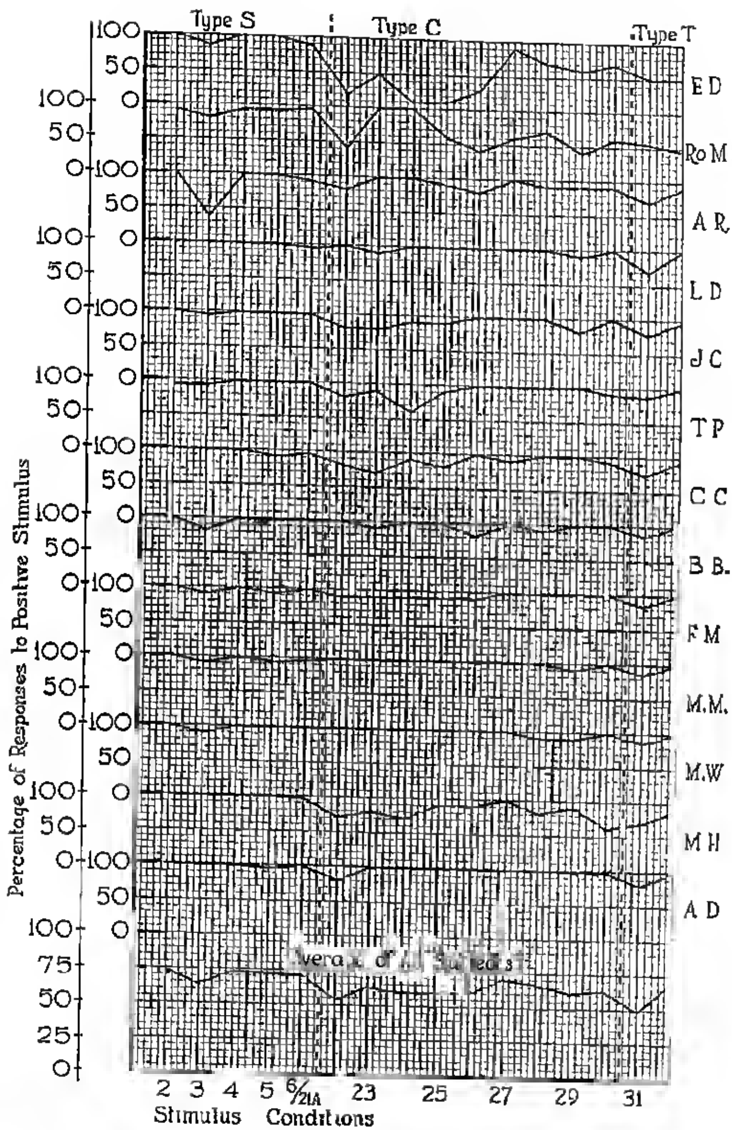

FIGURE 4

Percentage of Responses to the Positive Stimulus

tion 32 of Type T. This same order also holds for the majority of the individual children. The drop in the percentage of responses to the positive stimulus with the introduction of stimuli of Types C and T is interesting to note.

One other interesting trend is indicated in Table 5: within the group of Type C stimuli there seems to be a gradual rise in the percentage of responses to the positive stimulus which may indicate an adjustment period. At first the behavior which has been successful with stimuli on Type S may not appear to the child to be applicable to this new and different situation, but after reacting to several pairs of these objects the subject appears to become adjusted to this change and recognizes the positive stimulus almost at once. This trend is, however, not consistent as will be seen from a careful study of the table.

At first it might appear that there is a similar factor operating in the group of Type T stimuli when both Conditions 31 and 32 are considered together. This is, however, not the case. No such trend of improvement shows up within Condition 31 although it consists of 78 trials. The rise in percentage occurs with the change from Condition 31 to 32. Since the same figures were presented in both conditions, the manner of presentation, which varied, must be the factor accounting for the sudden rise. It will be recalled that in Condition 31 the figures were interpaired with each other; e.g., a four-sided figure was paired once with each of the other figures. Thus a circle was present in only 12 (15 per cent) of the 78 pairs, whereas in Condition 32, a circle was always present. Consequently, in the majority of the cases of Condition 31 the child's task was to pick the figure most nearly round, or the one with the larger number of sides. The mental set is certainly different in this situation. Since Table 5 shows that when the many-sided figures are paired with only a circle the response is consistently to the positive stimulus, the drop in the percentage of responses to the positive stimulus in Condition 31 can reasonably be attributed to the difference in the procedure and not to the figures themselves.

A few of the subject's verbal responses will be given. They clearly indicate that the attribute of roundness was consciously reacted to by some of the children. When the cylindrical objects were presented, L. D. called the positive stimuli "ball blocks," whereas spherical objects had been consistently called "balls." This

subject said the "ball stands" when the two-dimensional figures were introduced. Another child (*T.P.*) pointed out that "candy comes out where block is—a round block" (Condition 26) During the presentation of the two-dimensional figure he made such remarks as the following *"Where's round"? "Which is round"? "I look to see which is round one"* The following remark was made by *A. D* during Condition 2· *"If no ball in anyone* (compartment), *then I can't push any* (glass)*"* A little later, during Condition 5, he said, *"Everything got to be round to get candy out."*

G. Discussion of Results

Before beginning a discussion of the results a problem relating to procedure will be considered In several investigations of concept formation in the animal field the critical trial technique suggested by Kohler has been employed (5) The general procedure followed is to train the animal to the desired stimulus object, and then to present him with two new stimuli both of which are rewarded Since motivation is thereby equalized, it is assumed that any preference for one of the new stimuli will be the result of the previous training experience (cf, 6, p 128). The advisability of using this technique with children was investigated in a preliminary experiment It was observed that a child could be expected to make a response to a negative stimulus every once in a while These sporadic responses may be due to a temporary loss of interest, a decrease in motivation, or pure curiosity. The occurrence of such responses in a critical trial series changes the whole situation The child discovers that a response to the negative stimulus is rewarded This is usually a surprise to him and the novelty of obtaining candy from a different object seems to furnish an additional incentive to respond to it Thus the child may respond to this stimulus for several trials Eventually he may switch back to the positive stimulus, but in the meantime the percentage of responses to the positive stimulus has been made meaningless Alternating types of behavior may be the result of receiving candy from the negative stimulus On the next trial or two the child may return to the positive stimulus and then later go back to the negative The knowledge that candy is received from either stimulus may result in (*a*) a position habit, (*b*) a tendency to respond to both compartments, practically simultaneously, or (*c*) a random response to one

compartment and then another without any regard to the objects in the compartment

It seems to the writer, therefore, that the introduction of critical trials will eventually, but inevitably, lead to a breakdown of the original training If, on the other hand, critical trials are not presented, then the child soon learns that one stimulus will be rewarded and the other will not. Thus when a new pair of stimuli is introduced he knows that only one of them is associated with a reward If on the first trial he receives a reward, then the problem has been solved for his only task is to continue responding to this stimulus. If the negative stimulus was chosen on the first trial, then the child has only to remember that on all succeeding trials with this pair of objects he must respond to the other object. This line of reasoning leads to the conclusion that the method of scoring employed in transposition experiments is more valid, since the emphasis is placed on the first response (or first few responses). The main objection to such a scoring procedure is that random or chance responses, which are to be expected occasionally when young children are used as subjects, cause the reliability of single measures to be questioned Consequently, a middle course has been adopted Critical trials have not been presented in the main experiment reported herein, but instead the data have been analyzed to show the number of responses to the positive stimulus on the first trial of any condition as well as to all of the trials grouped

A group analysis of the data from the point of view of the number of responses to the positive stimulus (Table 3) indicates that in 85 per cent of the conditions the positive stimulus was chosen on the average not less than nine times out of ten A similar analysis from the point of view of the number of responses to the positive stimulus on the first trial shows that this stimulus was selected in 87 per cent of the conditions (Table 4) Thus the results from the two analyses agree and seem to indicate that for the group as a whole the concept of roundness was established A consideration of individual subjects shows that all consistently gave responses which indicated that they were reacting to the concept; yet some subjects with certain types of stimuli made such a low percentage of responses to the positive stimulus that it seems doubtful whether with these particular stimuli the concept operated. This was especially true of *E D* and *Ro M*, whose results are little better

than chance for all two-dimensional figures and for practically all cylindrical objects. Other instances of low percentages are noticed, but these are scattered among the subjects It may, therefore, be said that the remaining 11 subjects established the concept of roundness, but with certain stimulus conditions this response was weakened or broken down. An analysis of the first response to different conditions leads to similar conclusions, but the breakdown of Subjects *E. D.* and *Ro M* is not so clearly indicated A consideration of the responses to only the Type *C* stimuli does, however, indicate that the results of *E D.* and *Ro M* are comparable to chance expectation The results point to the conclusion that the concept of roundness was established for all subjects, but that the effect produced by varying the stimulus conditions differed from subject to subject

These variations in the stimulus conditions consisted of changes from spherical objects (Type *S*), to cylindrical objects (Type *C*), and finally to two-dimensional objects (Type *T*) The results of each group will be discussed first and will be followed by a discussion of the relations between the groups Each group will be considered as a unit and no attempt will be made to analyze the individual conditions within a group since the data on the variations (e g, background and the degree of roundness between objects of a pair) that occurred from condition to condition are too meager to warrant such treatment

The percentage of responses to the positive stimuli of Type *S* are all above 90 (Table 5), except in five instances Table 4 shows that six children made no first responses to negative stimuli of Type *S*: the largest number of such responses was three (*C G*). Thus there seems to be no doubt that *all* children responded on the basis of the concept of roundness.

With the introduction of the cylindrical objects (Type *C*) there is a definite drop in the percentage of responses to the positive stimulus When the curve for the percentage of responses to all positive stimuli of Type *C* (Figure 3) is inspected, it is observed that the initial drop is followed by a slight tendency to rise. This trend, although not present to the same degree in all subjects, suggests that the shift from spherical to cylindrical objects caused some readjustment This adjustment probably consisted of a shift in mental set from the expectation of spherical objects to the expecta-

tion of cylindrical objects as the rewarded stimulus. The established response was upset in only two subjects (*E. D.* and *Ro. M.*) The percentage of responses for the Type *C* stimuli of these two children approximates chance expectation. This is true not only for the average of Type *C* stimuli, but also for many of the separate conditions of this group of stimuli. Consequently the responses of these two children were not determined on the basis of roundness.

A considerable amount of variation in the percentage of the responses to the positive stimulus is caused by the introduction of two-dimensional figures (Type *T*). The results of *E. D.* and *Ro. M.* are chance scores; the other subjects' percentages are, however, greater than could be expected on the basis of chance alone. Consequently the other subjects responded on the basis of roundness, but with varying consistency. It will be recalled that the percentages for Condition 31 are consistently lower than for Condition 32. (Exceptions are *E. D.* and *Ro. M.*; their data will be omitted from the following discussion of Type *T*.) This has been attributed to differences in the method of presentation (i.e., in Condition 31 the comparison was between many-sided figures and many-sided figures or circle, while in Condition 32 one comparison figure was always a circle).

The high percentage of Condition 32 is taken to mean that the subjects' concept of roundness included two-dimensional roundness as well as three. Thus the interesting points brought out by the percentages of Conditions 31 and 32 are (*a*) that the child's generalization spreads to figures practically lacking a third dimension and (*b*) that he also tends to choose the more nearly round figure, but with less consistency than when choosing between round and non-round figures. This decrease in consistency may be due to a restricted concept of roundness, or the situation may be complicated by a difficult perceptual discrimination. That is, the child, as a result of reenforcements on previous trials, may be aware that to get candy the object most nearly round must be selected, but even though attempting to do this he may misjudge the correct stimulus.

When the relationship between the various types of stimulus is considered it is found that, following the preliminary training period with spherical objects, the response to the positive stimulus is most consistently made to Type *S* stimuli, less to Type *C* and least in Condition 31 of Type *T*, while Condition 32 is comparable

in some instances to Type C percentages and in others to Type S. This may be interpreted to mean that the subjects, except *E. D* and *Ro. M*, after having learned to select spherical objects consistently (above 90 per cent of the time) selected cylindrical and two-dimensional objects without further instructions or training. Thus their concept of roundness included all three groups of objects. The spread to cylindrical objects was not, however, immediate; an adjustment period was necessary, the duration of which was probably longer than the experiment since the consistency of selecting the positive cylindrical stimulus never reached that of selecting the spherical stimulus There seems to have been less disturbance by the introduction of the two-dimensional figures than by that of the cylindrical objects, but it is impossible to be sure of this since the latter always preceded the former, furthermore, some experience with the two-dimensional figures was obtained in Condition 31, the effect of which on the responses made in Condition 32 cannot be determined

The concept of spherical roundness which was established embraced cylindrical and two-dimensional objects to about the same extent in 11 of the 13 subjects Furthermore, the concept also allowed for the selecting of the roundest of two-dimensional non-round figures, but this spread of the concept is not so complete as in the case of cylindrical or two-dimensional round objects Stated in another way, in 11 subjects the range of equivalence for the concept of spherical roundness included cylindrical objects, two-dimensional figures, and the roundest of two-dimensional non-round figures

When the percentage of responses to the positive stimulus of all conditions is considered, it is evident that the child responded on a relative rather than an absolute basis. In no instance do the results indicate that an absolute attribute of the positive (or negative) stimulus was the determining factor of the response This is clearly indicated by the percentages of Table 5 as well as the number of first responses to the negative stimulus (Table 4)

The response is not dependent upon the physical identity of the stimuli but rather upon the characteristic which distinguishes the positive stimulus from the negative This characteristic is abstracted from the stimulus conditions and cannot be described in terms of quantitative physical attributes It is built up gradually by the

isolation of various differences between two stimuli. A final generalization is formed with the isolation of a characteristic which enables the correct stimulus to be chosen consistently.

H. Summary and Conclusion

The present investigation was concerned with a study of the concept of roundness in young children. Thirteen children, varying from three years of age to six years, three months, were trained to respond to a ball when paired with a block in a discrimination apparatus. The response consisted of a simple manual act. Only responses to the positive stimulus were rewarded.

The objects were exposed to the subjects in pairs. Three groups of positive stimuli were presented: spherical objects (Type *S*); cylindrical objects (Type *C*), and two-dimensional figures (Type *T*). Each consisted of nine or more different pairs of stimuli. In one condition (31) non-round figures were paired with non-round as well as round figures. It was found that:

1. All 13 subjects gave evidence of having established a concept of spherical roundness on the basis of the percentage of responses to the positive stimulus as well as on the basis of the number of responses to the positive stimulus on the first presentation of each pair of new and different stimuli.

2. This concept of spherical roundness was broad enough to include cylindrical objects as well as spherical objects in 11 of the 13 subjects.

3. The concept of spherical-circular roundness of 11 children included two-dimensional roundness, and the roundest figure of a pair of non-round two-dimensional figures.

4. The spread was not, however, equal for cylindrical objects, two-dimensional round figures, and roundest of two-dimensional figures. Cylindrical objects and round two-dimensional figures were most readily included in the pre-established concept, but even in these two instances a distinct drop in the percentage of responses to the positive stimulus is noted when these types of stimuli were first introduced. A more marked drop is found when non-round figures are paired. This may indicate that the limits of the range of equivalence of the concept of roundness were being approached, or that the results may be adversely influenced by the difficult perceptual discrimination that was involved.

In conclusion, the concept of spherical roundness established in the subjects included cylindrical roundness, two-dimensional roundness, and two-dimensional roundest. All responses were found to be relative in the sense that no single physical characteristic determined the response.[6]

References

1. Gellermann, L. W. Form discrimination in chimpanzees and two-year-old children. I. Form (triangularity) *per se*. *J. Genet. Psychol.*, 1933, 42, 3-27.
2. ——— Chance order of alternating stimuli in visual discrimination experiments. *J. Genet. Psychol.*, 1933, 42, 206-208.
3. Hicks, J. A., & Stewart, F. D. The learning of abstract concepts of size. *Child Devel.*, 1930, 1, 195-203.
4. Klüver, H. Behavior Mechanisms in Monkeys. Chicago: Univ. Chicago Press, 1933. Pp. xiv+387.
5. Kohler, W. Simple structural functions in the chimpanzee and in the chicken. *In* (Ellis, W. D.) *A Source Book of Gestalt Psychology*. New York: Harcourt, 1938. Pp. xiv+403.
6. Lashley, K. S. The mechanism of vision. XV. Preliminary studies of the rat's capacity for detail vision. *J. Genet. Psychol.*, 1938, 18, 123-193.
7. Munn, N. L., & Stiening, B. R. The relative efficacy of form and background in a child's discrimination of visual patterns. *J. Genet. Psychol.*, 1931, 39, 73-90.
8. Roberts, Katherine R. The ability of preschool children to solve problems in which a simple principle of relationship is kept constant. *J. Genet. Psychol.*, 1932, 40, 118-135.
9. Thrum, Martha E. The development of concepts of magnitude. *Child Devel.*, 1935, 6, 120-140.
10. Warden, C. J., & Rowley, J. The discrimination of absolute vs. relative size in the ring dove. *J. Comp. Psychol.*, 1929, 9, 317-337.
11. Welch, L. The development of discriminaton of form and area. *J. of Psychol.*, 1939, 7, 37-54.

Normal Child Development Study
Babies Hospital
630 West 168 Street
New York City

[6]The writer would like to express his appreciation to R. S. Childs for his assistance in planning and constructing the apparatus.

The Journal of Genetic Psychology, 1940, 57, 317-326

THE RIGHTS-MINUS-WRONGS METHOD OF CORRECTING CHANCE FACTORS IN THE TRUE-FALSE EXAMINATION*

Gill, Colorado

CLARENCE CARL MOORE

"Children's attitudes are largely affected by a concept which they call fairness. . The fair teacher is one who is accurate" (2)

Fairness does not consist of measurement which purports to measure exactly and only gives a rough estimate The mere fact that a test gives the same results when scored by two or more people is no assurance that it is fair to the pupils who are measured Objectivity in scoring is not necessarily an assurance of an accurate measuring instrument

Tests are used for determining school marks and school marks are used for a great many purposes some of which determine the pupil's attitude toward school and life For most of these purposes marks obtained from alternate-response tests are generally assumed to be accurate measures of a pupil's status.

Most studies that have been made for the purpose of measuring the amount of error in the true-false test have failed to deal with a small number of chance units such as are introduced when a group of 25, 60, or 100 pupils are measured Instead of a number of units comparable to that of an ordinary true-false test, from 500 to 50,000 units have been used There is a vast difference in the amount of error obtained from guessing at 25 items taken 200 times and that obtained from guessing at 5,000 items Another oversight of some writers is to distinguish between the number of items on a test and the number of guesses the subject makes in answering the test A student who is ready for an examination does not guess at all of the items which a test contains. It is also uncommon for true-false tests of more than 200 items to be administered under ordinary testing conditions.

The purpose of this study was to create a testing situation in which the subjects guessed at a certain number of statements, to score these guesses by means of a true-false scoring key, and to

*Received in the Editorial Office on August 22, 1939

determine the extent of agreement or departure of the subjects' guesses with that which might be expected from the use of the rights-minus-wrongs formula.

In order to make sure that the tests provided only for pure guesses the following procedure was followed: (*a*) The tests were written in Latin, (*b*) Latin words within each supposed sentence or item were arranged in a random fashion (they probably would not have made sense to a person who could translate Latin), and (*c*) the tests were administered to students who had never studied Latin.

The tests were mimeographed in lengths of 10, 25, 40, 70, and 100 items. Each length of test was administered to groups of 25, 60, and 100 pupils, respectively. Tables were built from the results. There was a frequency distribution for each of the different lengths of tests as administered to 25 pupils in each group, 60 pupils in each group, and 100 pupils in each group. Since some of the results showed an excess for right responses and some of the results showed an excess of wrong responses, a constant was added to each score. This constant in all cases equalled 50.

The reason for adding a constant was explained in a previous article (1). The same tabular arrangement was followed in this study as that shown in the preliminary study with the exception that the percentile points of each distribution has been computed. Since the procedure in rearranging the tables to care for the constant should be evident from the preliminary study, the raw-data tables have been omitted from this study. The terminology applied to the distributions of this study differs slightly from that applied to the distributions of the first study. In the preliminary study the comparable distributions were called "Hypothetical"; in this study they are called "Constant Added." The term "Constant Added" is reduced to the symbol *"CA."* The latter title is probably more characteristic of the distributions than the former.

The Constant Added Chance Distributions

Table 1 shows five *"CA"* chance distributions of Latin test scores based upon the assumption that classes with 25 pupils in each class take five-true-false Latin tests each of which is 150, 120, 90, 75, and 60 items in length, that on each of the different lengths of tests each pupil knows exactly 50 items, and that each pupil

TABLE 1

FIVE "CA" CHANCE DISTRIBUTIONS OF LATIN TEST SCORES BASED UPON THE ASSUMPTION THAT CLASSES WITH 25 PUPILS IN EACH CLASS TAKE FIVE TRUE-FALSE LATIN TESTS EACH OF WHICH IS 150, 120, 90, 75, AND 60 ITEMS IN LENGTH, THAT ON EACH OF THE DIFFERENT LENGTHS OF TESTS EACH PUPIL KNOWS EXACTLY 50 ITEMS, AND THAT EACH PUPIL GUESSES AT THE REMAINING 100, 70, 40, 25 AND 10 ITEMS ACCORDING TO EACH OF THE RESPECTIVE LENGTHS OF THE TESTS MENTIONED ABOVE, 1939

Score	Frequencies of the following number of guesses				
	100	70	40	25	10
1	2	3	4	5	6
70	1	1			
62		2			
60		1	2		
58	2	3	2		
57				1	
56	2	3	2		
54	6	1	2		4
53				6	
52		3	4		4
51				6	
50	2	3	2		6
49				4	
48	2	3	6		5
47				3	
46	1				6
45				5	
44	2	1	2		
42	1		1		
40	2	3	2		
38	2	1			
36	1				
28	1				
Total	25	25	25	25	25
Guesses	2,500	1,750	1,000	625	250
$P_{100}—P_0$	42.00	32.00	20.00	12.00	8.00
$P_{90}—P_{10}$	20.00	21.50	16.50	8.50	7.92
$P_{75}—P_{25}$	12.38	9.67	7.33	5.42	4.78

Total number of guessed items in this table, 6,125. Total excess of right guesses over wrong guesses, 12

guesses at the remaining 100, 70, 40, 25, and 10 items according to each of the respective lengths of the tests mentioned above Column 1 shows the scores parallel to which are the frequencies shown in Columns 2, 3, 4, 5, and 6 Column 2 shows, that for the 25 tests containing 100 items, there was one paper with a score of 70, two papers with scores of 58, two papers with scores of 56, six papers with scores of 54, and so on through the column

The percentile points at the bottom of the column show that when a group of 25 pupils each guess at 100 items of a true-false test containing 150 items the luckiest pupil may have an advantage of 42 items over the unluckiest pupil, the luckiest 10 per cent of the group may have an advantage of 20 items over the unluckiest 10 per cent of the group, and the luckiest 25 per cent of the group may have an advantage of at least 12 items over the unluckiest 25 per cent of the group

Column 5 shows that when a group of 25 pupils each guess at 25 items of a true-false test containing 75 items the luckiest pupil may have an advantage of at least 12 items over the unluckiest pupil, the luckiest 10 per cent of the group may have an advantage of at least 9 items over the unluckiest 10 per cent of the group, and the luckiest 25 per cent of the group may have an advantage of at least 5 items over the unluckiest 25 per cent of the group.

Column 6 shows that when a group of 25 pupils each guess at 10 items of a true-false test containing 60 items the luckiest pupil may have an advantage of at least 8 items over the unluckiest pupil of the group, the luckiest 10 per cent of the group may have an advantage of at least 8 items over the unluckiest 10 per cent of the group, and the luckiest 25 per cent of the group may have an advantage of at least 5 items over the unluckiest 25 per cent of the group

The percentile points of Columns 3 and 4 may be interpreted in the same manner as Columns 2, 5, and 6 have been interpreted

Table 2 shows five *"CA"* chance distributions of scores based upon the assumption that classes with 60 pupils in each class take five true-false Latin tests each of which is 150, 120, 90, 75, and 60 items in length, that on each of the different lengths of tests each pupil knows exactly 50 items, and that each pupil guesses at the remaining 100, 70, 40, 25, and 10 items according to each of the respective lengths of the tests mentioned above.

Column 1 shows that when a group of 60 pupils each guess at 100 items of a true-false test containing 150 items the following advantages may occur due to mere chance the luckiest pupil over the unluckiest pupil, 36 items, the luckiest 10 per cent over the unluckiest 10 per cent, 24 items, the luckiest 25 per cent over the unluckiest 25 per cent, 14 items.

Column 6 shows that when a group of 60 pupils each guess at 10 items of a true-false test containing 60 items the following advantages

TABLE 2

FIVE "*CA*" CHANCE DISTRIBUTIONS OF SCORES BASED UPON THE ASSUMPTION THAT CLASSES WITH 60 PUPILS IN EACH CLASS TAKE FIVE TRUE-FALSE LATIN TESTS EACH OF WHICH IS 150, 120, 90, 75, AND 60 ITEMS IN LENGTH, THAT ON EACH OF THE DIFFERENT LENGTHS OF TESTS EACH PUPIL KNOWS EXACTLY 50 ITEMS, AND THAT EACH PUPIL GUESSES AT THE REMAINING 100, 70, 40, 25 AND 10 ITEMS ACCORDING TO EACH OF THE RESPECTIVE LENGTHS OF THE TESTS MENTIONED ABOVE, 1939

Score	Frequencies of the following number of guesses				
	100	70	40	25	10
1	2	3	4	5	6
68		2			
64	1	1	2		
63				1	
62	2	1	1		
60	2	11	4		
59				2	
58	3	3	2		1
56	3	4	9		1
55				4	
54	2		7		6
53				9	
52	5	5	8		10
51				9	
50	5	8	5		13
49				9	
48	3	9	6		16
47				12	
46	3	4	5		8
45				3	
44	7	8	4		5
43				9	
42	6		3		
41				1	
40	4	1	3		
38	4	1			
37				1	
36	2	1			
34	4				
32	2		1		
28	2	1			
Total tests	60	60	60	60	60
Guesses	6,000	4,200	2,400	1,500	600
$P_{100}—P_0$	36.00	40.00	32.00	26.00	14.00
$P_{90}—P_{10}$	24.33	16.89	17.17	11.61	8.54
$P_{75}—P_{25}$	14.00	12.34	10.34	7.40	5.07

Total number of guessed items in this table, 14,700. Total excess wrong guesses over right guesses, 190.

may occur due to mere chance the luckiest pupil over the unluckiest pupil, 14 items, the luckiest 10 per cent over the unluckiest 10 per cent, 9 items; the luckiest 25 per cent over the unluckiest 25 per cent, 5 items

The non-uniformity with which factors of chance may operate is indicated by the advantage of the luckiest pupil over the unluckiest pupil as shown in Columns 2 and 3 In Column 2 this advantage for 100 guesses amounted to 36 items and the same advantages as shown in Column 3 amounts to 40 items.

Table 3 shows five *"Cd"* chance distributions of scores based upon the assumption that classes with 100 pupils in each class take five true-false Latin tests each of which is 150, 120, 90, 75, and 60 items in length, that on each of the different lengths of tests each pupil knows exactly 50 items, and that each pupil guesses at the remaining 100, 70, 40, 25, and 10 items according to each of the respective lengths of the tests mentioned above.

Column 2 shows that when a group of 100 pupils each guess at 100 items of a true-false test containing 150 items the following advantages may occur due to mere chance: the luckiest pupil over the unluckiest pupil, 44 items, the luckiest 10 per cent over the unluckiest 10 per cent, 24 items; the luckiest 25 per cent over the unluckiest 25 per cent, 15 items

Column 6 shows that when a group of 100 pupils each guess at 10 items of a true-false test containing 60 items the following advantages may occur due to mere chance: the luckiest pupil over the unluckiest pupil, 14 items, the luckiest 10 per cent of the group over the unluckiest 10 per cent of the group, 7 items, and the luckiest 25 per cent of the group over the unluckiest 25 per cent of the group, 4 items

Conclusions

This study indicates that for tests having a number of items equal to 10, 25, 40, 70, and 100 at which the pupils must guess the rights-minus-wrongs method of scoring does not reduce the influence of the factor of chance sufficiently to make the true-false test a reliable measuring instrument unless the proportion of guessed items to known items is comparatively small.

The greater the number of items on a test at which the pupil must guess the greater are his chances of making an occasional high or low score through chance alone.

TABLE 3

FIVE "*CA*" CHANCE DISTRIBUTIONS OF SCORES BASED UPON THE ASSUMPTION THAT CLASSES WITH 100 PUPILS IN EACH CLASS TAKE FIVE TRUE-FALSE LATIN TESTS EACH OF WHICH IS 150, 120, 90, 75, AND 60 ITEMS IN LENGTH, THAT ON EACH OF THE DIFFERENT LENGTHS OF TESTS EACH PUPIL KNOWS EXACTLY 50 ITEMS, AND THAT EACH PUPIL GUESSES AT THE REMAINING 100, 70, 40, 25 AND 10 ITEMS ACCORDING TO EACH OF THE RESPECTIVE LENGTHS OF THE TESTS MENTIONED ABOVE, 1939

Score	Frequencies of the following number of guesses				
	100	70	40	25	10
1	2	3	4	5	6
70	1	1			
68	2	1			
66		1	1		
64	1	2	2		
63				1	
62	2	5	1		
60	2	11	6		
59				2	
58	7	7	6		1
57				2	
56	7	8	14		1
55				5	
54	9	2	11		13
53				18	
52	6	9	14		18
51				17	
50	7	14	9		25
49				18	
48	5	14	13		22
47				16	
46	4	4	5		15
45				8	
44	9	9	9		5
43				10	
42	7	1	3		
41				2	
40	8	6	5		
38	8	2			
37				1	
36	3	1			
34	6				
32	2	1	1		
28	3	1			
26	1				
Total tests	100	100	100	100	100
Guesses	10,000	7,000	4,000	2,500	1,000
$P_{100}—P_0$	44 00	42 00	34 00	26.00	14 00
$P_{90}—P_{10}$	24 00	22 33	17 78	12 40	7.44
$P_{75}—P_{25}$	14 83	10.71	8 40	5 83	4 44

Total number of guessed items in this table, 24,500 Total excess of wrong guesses over right guesses, 106

It is possible for a test of 50 items to be more valid and reliable than another test of 100 items This depends upon the number of items at which the pupil guesses in proportion to the total number of items which the test contains

Some writers have recommended that subjects be instructed not to guess If a student guesses, he must suffer the consequences when he makes unlucky guesses He also profits when he makes lucky guesses Then too, there is the person who is amenable to instruction and who may leave out some of the statements of which he is not sure because he does not know for sure if he knows or does not know In other words, there is no scientific way of finding out whether the subject guesses or does not guess, whether he pats himself on the back when he is lucky or whether he blames the teacher when he is unlucky However, there has been a slight advantage found in favor of accuracy when pupils are instructed not to guess

The writer recommends the scoring of doubtful items before the individual score of class members is computed Those items which fail to be answered correctly by from 55 to 95 per cent of the class should be eliminated before the score of the individual members is determined Since approximately 50 per cent of a group answering an examination statement will answer it correctly and 50 per cent incorrectly as a result of chance, when chance operates according to expectation, a margin of 5 per cent is not too much to allow for the variation in the uniform operation of it. Less than 50 per cent of a class answering a statement correctly indicates results which are negative to the results expected by the teacher when the scoring key was constructed Such a situation may indicate that the teaching was negative to the concept which the teacher had in mind, that there was a high degree of chance involved, or that the teacher's viewpoint is wrong or at least different from the viewpoint of the class concerning this particular statement

In the first and last situations, the teacher has not produced enough change in the pupil's way of thinking to make the statement worthy of inclusion in the test In any event, it is better to eliminate the item, for it reflects either on the teacher's judgment of how well he has taught, or on his ability to teach in a comprehensible manner

Even when true-false tests have been constructed so that the difficulty of the test is in the proper proportion to the accomplishment of the class, they are still only rough estimates of the individual

achievement of certain pupils within the group If care is not exercised in the wording of the statements and in the amount of difficulty the test contains, alternate-response tests should not be used at all

Those who continue to measure with alternate-response tests should understand their limitations as exact instruments of measurement The writer obtained an actual class distribution of scores which had been constructed from the scores of 60 pupils This distribution had a range of from 55 to 90 with a median of 73 9 It was almost a normal distribution The scores of this distribution had been divided into five groups by means of the 1 6 *Q-Md* method (3) The purpose of this distribution was to determine the grades or marks which the individual members of a class should receive in connection with the *A*, *B*, *C*, *D*, and *F* plan of awarding class marks In order to estimate the advantage and disadvantage due to guessing at 25 items of a 100-item test, the writer compared the results of Table 2, Column 5, with the above distribution If chance operated in this instance as it did in Table 2, Column 5, if the distribution in which the pupil should be placed each testing period were a normal distribution, and if the pupil knew exactly 75 of 100 items and guessed at the remaining 25 items on a series of 60 tests, his marks would be approximately as follows one mark of *A*, 14 marks of *B*, 30 marks of *C*, 14 marks of *D*, and one mark of *F* It should be emphasized that the pupil was assumed to know an equal number of statements each testing period

There may be some who will conclude that such a distribution of such a pupil's marks is immaterial due to the fact that he is an average pupil and that the final outcome of his 60 marks gives him an average standing of *C* That is, the final summation of the results of many examinations gives the pupil approximately his true status The writer can find little fault with such a philosophy as it applies to many tests It is basically sound However, the psychological effect of such inconsistency may lead to immediate attitudes on the part of the pupil which are basically unsound This may be especially true of the immediate effect upon the attitudes of younger pupils

As an illustration, suppose that an average pupil gets a mark of *D* at one testing period and a mark of *B* at the next testing period, or that a pupil should work very diligently previous to one testing period and that he is so unfortunate with his guesses that he receives

an average mark of *C* and that he fails to work up to his ability previous to the succeeding testing period but is so lucky with his guesses that he receives the same mark. Isn't such a situation as either of these likely to lead the pupil to assume a passive attitude toward his school program?

Teachers, who feel that the true-false examination is an exact instrument of measurement, may also be led to wrong conclusions concerning the child's efforts. Parents to whom the reports from such measurements are sent may penalize the child for conditions which are entirely beyond his control. If there are situations in which the alternate-response test measures certain knowledge, attitudes or skills better than other types of tests, they should be used But why not accept a scientific attitude toward such testing situations and frankly admit that their accuracy is questionable even though they are carefully constructed and properly applied.

References

1. MOORE, C C Factors of chance in the true-false examination *J. Genet. Psychol.*, 1938, 53, 215-229
2. ——— The rights-minus-wrongs method of correcting chance factors in the true-false examination Unpublished Doctor's Field Study, Colorado State College of Education, Greeley, Colorado
3. RINSLAND, H D Constructing Tests and Grading New York, Prentice-Hall, 1937

Gill, Colorado

The Journal of Genetic Psychology, 1940, 57, 327-344

NERVOUS AND HORMONAL FACTORS IN THE MATERNAL BEHAVIOR OF THE MOUSE*[1]

Department of Anatomy, Yale University School of Medicine, and Laboratoire de Synthèse Atomique de Paris

CHARLES PHILIPPE LEBLOND

Maternal behavior comprises the activities connected with the care of offspring. It is ordinarily seen during lactation, i.e., the period between the birth of the young and the time of their weaning. Since one of the main features of this period is the release by the hypophysis of the hormone, prolactin, which determines milk secretion (19), attempts have been made to induce maternal behavior with either crude hypophyseal extracts (7, 34, 18) or with prolactin (26, 27, 28, 21).

Apart from the hormonal factors, the activities involved in maternal behavior imply nervous mechanisms which determine its pattern. In fact, lesions of various parts of the nervous system have been reported to interfere with this behavior (29, 22, 2).

The present investigation sought to determine whether or not hormones are necessary to activate the nervous mechanisms of maternal behavior in the mouse.

A. MATERIAL

Mice of pure strains (mainly *A*, also *CH1*, *C3H*, *C*, *JK*, *CBAN*, *F*, *L*, *N*) were used in this investigation. The tests were performed in metallic cages (12" x 6" x 3") containing shavings and a little cotton. A wire cover closed the cage, the inside of which was entirely visible.

Usually one animal was kept in each cage. The mice were used to being looked at and handled. Most of the time they did not attempt to escape the author's hand, neither did they attack it.

The animals were given free access to water and purina fox chow.[2]

*Received in the Editorial Office on August 31, 1939.

[1]A preliminary report of the results of this investigation has been published elsewhere (14).

[2]We are indebted to Dr. L. C. Strong for the mice.

B Behavior During Lactation

The understanding of the experiments reported below necessitates a brief account of the behavior of lactating mice

The maternal activities of the mouse are very similar to those which have been described for the rat (33, 20, 34) The maternal mouse, usually after sniffing at the young put in her cage, retrieved it, holding it between her teeth by the skin of the back, and carried it to the nest After depositing the young, she went back to the place of retrieving and continued to pick up young and carry them to the nest as long as she found one In one instance, 50 newborn mice put in the cage of a lactating mouse were readily retrieved It was not possible to observe a discrimination between foster young and the own young of the animal

In the nest, the lactating mouse licked the young thoroughly, frequently confining her licking to the external genitalia She then huddled over the young When the young were sucking the nipples, she seemed to be in a relaxed position, legs stretched over the young, remaining immobile in spite of the crawling and shifting of the nurslings This calmness was in contrast to the habitual restlessness of mice, and was especially noticeable when a cage containing several mice was opened, usually all the animals left the nest except the females in the act of nursing

During the lactation period, various activities were observed which, unlike retrieving, licking, and cuddling, were not directly provoked by the presence of young (*a*) Nest building was characterized by digging and carrying material to the site of the nest to be constructed, resulting in a hollow nest frequently enclosing the animal (*b*) Defense reactions were observed mostly in isolated animals. Such reactions could be seen when a mouse was introduced into a cage where a lactating animal was kept with her young the newcomer was usually fought When young were born in a cage where several mice were present, it was sometimes observed that after the parturition the mother did not permit the other adult mice to penetrate into the common nest.

C Methods and Results

Maternal behavior was examined by placing a young for five minutes in the cage of the animal to be tested. Killing of the young was only rarely observed, contrary to what happens in colonies of

mice which are not used to being handled Two types of tests were used: either a single newborn mouse was presented or mice pups of various ages were successively put in the cage

In both tests three types of reactions were noted (*a*) Retrieving was rated "one," when the young was picked up and dropped anywhere in the cage, "two," when the young was carried to the nest, and "three," when the mouse then came back to the spot where the pup had been found (*b*) The care of the young was rated "one" when licking of the young for over 60 seconds was observed, and "two" when the mouse would sit over the young in the nest until the end of the five-minute test (*c*) Nest building was sometimes found to occur It was rated "one" when there was digging in the shavings; "two" when material was carried to the nest, "three" was applied to the building of a hollow nest over the young. The arbitrary addition of the figures for retrieving, care, and nest building furnished a "maternal index" from 1 to 8 which appeared to be helpful in this investigation (13).

1 *Provocative Value of Young Mice*

The value of the tests for maternal behavior was found to depend on the age of the young used for testing We have frequently observed that a mouse will display full maternal activity towards a newborn mouse, but will pay little attention to a 10-day-old one, and will ignore one 15 days old. The ability of the young to elicit maternal activity appeared to decrease as the age increased, as has been reported for the rat (34) This effectiveness of the young to evoke expressions of maternal behavior is referred to in this paper as their "provocative value."

In order to measure the provocative value of young of different ages, the frequency of the retrieving of young of four different ages was determined. For this purpose 15 ($\pm$1)-day, 10 ($\pm$1)-day, 5 ($\pm$1)-day and new-born (0-1 day) young were successively placed for 5 minutes each in the cage of the same adult animal [3] One hundred and ninety-two such tests were performed at 2-3-day intervals, using 20 lactating females These tests showed that new-

[3]It is possible, especially in view of the "sensitization" factor discussed below, that failure to present young of different ages in a mixed, balanced order, may have affected the results slightly, tending to increase the number of younger, and to decrease the number of older, pups retrieved

TABLE 1
FREQUENCY OF THE RETRIEVING OF YOUNG OF DIFFERENT AGES

	Tests with lactating females		Tests with parental males	
	Number retrieved	Frequency of retrieving	Number retrieved	Frequency of retrieving
	$N = 192$		$N = 52$	
0-1-day-old young	160	83%	44	85%
5-day-old young	149	78%	41	79%
10-day-old young	104	54%	21	40%
15-day-old young	21	11%	8	15%

born and 5-day-old young were retrieved much more frequently than were older pups (Table 1, Columns 2 and 3) The frequencies of the retrieving of new born, 5-day, 10-day and 15-day young by lactating females were 83 per cent, 78 per cent, 54 per cent, and 11 per cent respectively. Using these figures as an index of the provocative values of the young, one sees that these values decrease relatively more rapidly as the age of the young increases (Figure 1).

Since a 15-day-old young has but little provocative value, one may assume that a mouse which retrieves it has a strong maternal drive On the other hand, a mouse which would respond only to a newborn young (strong stimulus) has only a weak maternal drive

If young of different ages are presented to a maternal mouse, the oldest young retrieved will indicate the smallest stimulus capable of eliciting maternal responses The age of the oldest young retrieved may be used as a figure for the intensity of the maternal behavior, as was done by Wiesner and Sheard for the rat. This convenient method was used in our investigation.

Instead of using the age of the oldest young retrieved as an index of maternal behavior, one might derive from the curve (Figure 1) the corresponding provocative values The figures, thus calculated, did not appreciably change the results given below and will not be reported

2. *Variations in Maternal Behavior During Lactation*

In most mammals, maternal behavior has close relationship to lactation, appearing and disappearing with it (15) Furthermore, during lactation variations in maternal behavior have been found In the rat, this behavior reached a peak soon after parturition and

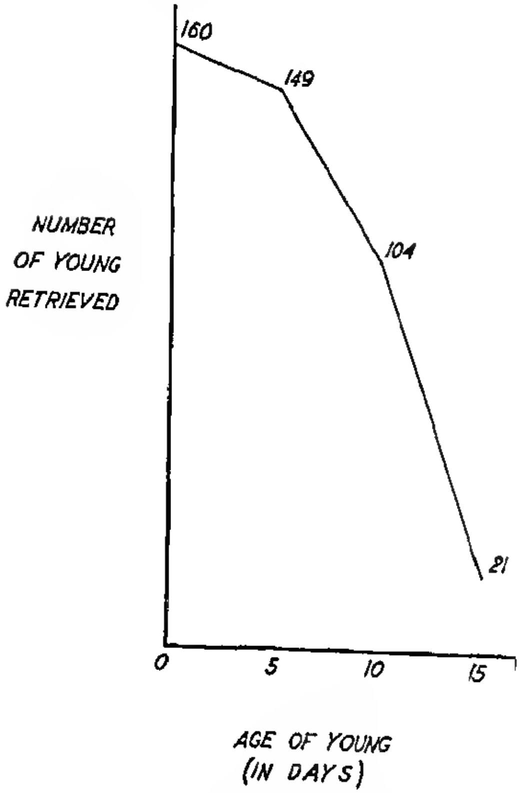

FIGURE 1

NUMBER OF YOUNG MICE OF DIFFERENT AGES RETRIEVED IN 192 TESTS MADE WITH LACTATING FEMALES

The youngest pups were the most frequently retrieved and, therefore, had the greatest provocative value

decreased slowly until the end of the lactation period (33, 34). Casual observations revealed that this might be the case for cats and rabbits also

Twenty mice were tested three times a week, starting a few days before parturition and continuing until after the young were weaned. The animals were separated from their young for a half hour before the test The age of the oldest young retrieved, deter-

332 JOURNAL OF GENETIC PSYCHOLOGY

TABLE 2

VARIATIONS IN MATERNAL BEHAVIOR DURING LACTATION

		Days before parturition		Days after parturition						Three weeks after end of lactation	Three months after end of lactation
		3 Days	1-2 Days	0-3 Days	4-7 Days	8-11 Days	12-15 Days	16-19 Days	20-23 Days		
Number of tests		11	17	33	35	30	30	30	9	19	20
Average maternal index		3	3 4	4 5	3 6	3 8	3 8	3 2	3 1	3 6	3 8
Average age of oldest young retrieved (in days)	Mouse	5 9	7 1	8 5	6 9	8 5	8 1	6 4	5 7	9 1	7 8
	Rat (after Wiesner & Sheard)	—	—	11 4	5 8	6 1	5 5	3 8	3 4	—	—

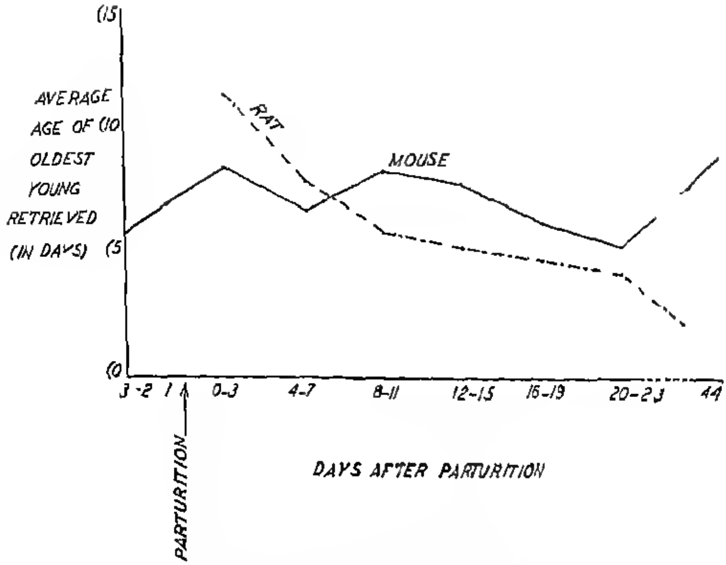

FIGURE 2

VARIATIONS OF MATERNAL BEHAVIOR INDICATED BY THE VARIATIONS OF THE AGE OF THE OLDEST YOUNG RETRIEVED, BEFORE, DURING AND AFTER LACTATION

mined as described in the preceding section, was used as a figure for the intensity of the maternal behavior The maternal index was calculated from the observations on the new born young presented at the end of the test In addition, tests on eight mice were made as late as three weeks, and on four mice, over three months after the end of lactation

The results (Table 2 and Figure 2) have been supplemented with data on the rat, which were calculated from Wiesner and Sheard's figures showing the peak of maternal behavior at parturition and its gradual decrease during lactation, as reported for this species (34, 20) In the mouse there were no such variations of maternal behavior Neither the maternal index nor the average age of the oldest young retrieved showed important modifications.

However, the figures suggested the possibility of a slight increase of maternal behavior before parturition and of a small peak just after parturition (Table 2) Analysis of the data showed that these

features existed only in primiparous females This was apparently due to hormonal shifts before and at parturition, although the stimulation by the young used for the tests might have been the agent of this increase, as indicated below in this paper. The increase of maternal behavior at the first parturition was more definite with the mice of the strains N and F (two mice of each) Furthermore, in these animals, the average age of the oldest young retrieved decreased during lactation, from 8 1 soon after parturition to three 10 days later, and 2 5, 21 days later The five multiparous females did not show variations of maternal behavior before, during, or after lactation, but the averages suggested a more intense behavior than in primiparous females (Table 3). The number of positive tests

TABLE 3
INTENSITY OF MATERNAL BEHAVIOR IN PRIMIPAROUS AND MULTIPAROUS FEMALES

	Number of animals	Number of tests Total	Positive	Per cent of positive tests	Maternal index	Average age of oldest young retrieved (in days)
Primiparous	15	132	114	76%	3.7	7 0
Multiparous	5	35	30	86%	3 8	9.0

was larger and the average age of the oldest young retrieved, greater in multiparous than in primiparous females However, the average maternal index was the same in the two cases

On the other hand, an intense and immediate maternal behavior was observed in six out of eight females kept away from young for three weeks and in four out of four females kept away from young for three months after the end of the lactation period (Table 2) One of these females, in addition to the usual test was given 50 new-born young. She retrieved, licked and cuddled them eagerly, and displayed definite unrest, when the young had been removed So typical a behavior outside lactation confirmed the small importance that the hormonal set-up of lactation had on the maintenance of maternal activities.

3 *Maternal Behavior in Various Physiological and Experimental Conditions*

A behavior similar to that shown by lactating and post-lactating

females could be, under definite conditions, observed in males and virgin females. Thus, when male mice were kept in the same cages as lactating females, we frequently noticed that the males would pick up the young, lick them and cuddle them in a way that could hardly be distinguished from that of the lactating females.

Methodical investigations showed males with a definite "maternal" behavior in all the strains of mice examined. Ninety-four per cent of 47 males from 22 to 120 days of age were found to exhibit maternal behavior. It was usually necessary to keep the animals with new-born young for periods varying from a few minutes to several days previous to the test. This treatment was prolonged until the animal would show maternal behavior, but usually not over four days. The effect of the young in such instances is referred to in this paper as "sensitization." A group of 13 parental sensitized males were tested with young of different ages (15-, 10-, 5-, and 1-day old) in the same way as the lactating females above described. The results (Table 1, Columns 4 and 5) show that the provocative value of the young is practically the same for parental males and lactating females.

After a "sensitization" as above, 12 out of 14 virgin females showed evidence of the presence of maternal behavior.

Then, tests were performed on mice deprived of sex hormones either by castration or by hypophysectomy. Two male animals were castrated at birth, and three after puberty. All but one of the animals (in the latter group) showed typical maternal behavior after sensitization.

Suppression of the hypophysis produces multiple hormonal deficiencies, since it is known that this gland is wholly responsible for the activity of the sex and mammary glands, and partly responsible for the functions of the thyroid and adrenal glands. When females were hypophysectomized during lactation, milk secretion was immediately and completely arrested in spite of the continued suckling by the young left in the cage. However, in these animals, retrieving, licking, and cuddling of the young continued as before the operation. Furthermore, when animals were hypophysectomized without having ever been in the presence of young, it was possible to induce them to be maternal through "sensitization" after the operation (13).

Eight out of nine mice,[4] having received 10,000 to 16,000 *R.U.* of oestrin in fractionated doses from birth until they were tested at the age of 121 to 203 days, became readily maternal in the presence of young.

A few observations on the maternal behavior of rats showed a fairly complete pattern in two out of three hypophysectomized lactating female rats (13) and in three out of 10 normal virgin female rats after sensitization. Retrieving alone was found in virgin females (5 out of 10), normal (4 out of 6) and castrated (2 out of 3) male rats five months of age, also after sensitization.

D. Discussion

1. *Physiological Phenomena of Lactation (Rat, Mouse)*

A brief account of the physiology of lactation is necessary here for the interpretation of the results reported above.

Milk secretion is initiated by the release of *prolactin* from the hypophysis as a consequence of the interruption, at parturition, of the hormonal secretions from the placenta and the corpus luteum of pregnancy (19). A reflex from the empty uterus also conditions milk secretion, since Selye has shown that lactation may be inhibited by filling the uterus with paraffin after the fetuses have been removed (30).

The maintenance of lactation, which lasts for about 21 days, is due to suction on the nipples, since the milk secretion stops when the young are prevented from sucking. By means of nervous pathways, the suction acts on the hypophysis, causing it to release prolactin, thus maintaining lactation (31, 10). Suction is also the cause of the anoestrus of lactation: no follicles ripen in the ovary and the animal does not come in heat. This too, may be due to the release of prolactin, inasmuch as the injection of this hormone prevents estrus (6, 12, 15).

The giving of milk to the young causes the mother to lose a considerable quantity of nutritive elements. Some rats have been found to secrete over night from 3 to 8 cc. of a very rich milk (protein 11.70 per cent, fat 14.79 per cent) (5). As is expected, the appetite is increased, an average intake of three times the normal amount of fat and protein is observed (24). The ensuing

[4]We are indebted to Dr. W. U. Gardner for these animals.

higher metabolism may account for the rise in average temperature to 38.4° C in these lactating rats, instead of the 37.6° C of the controls. This supernormal temperature, regularly found only in animals with over four young is most unstable, frequently going over 39° C (16). Probably as another consequence of the increased metabolism, the activity of the lactating animal, as measured in a revolving drum, is very much decreased (32).

2. *Causes of the Main Forms of Maternal Behavior*

In opposition to such activities as nest building and defense reaction which, as a rule, were not provoked by the presence of young, the main forms of maternal behavior were defined as those which a young regularly elicited in a parental animal, namely retrieving, licking, and cuddling. It should be emphasized that only when the latter activities occurred in closely associated succession did they constitute true maternal behavior. Licking for instance, was not necessarily connected with maternal behavior, as it was observed when a strange adult mouse was introduced into the cage. Retrieving alone was similar to nest building. Richter is even of the opinion that rat pups constitute good nest building material. Retrieving being thus explained on the basis of the nest building drive, the maternal behavior might be reduced to that drive (25). The following reasons may be given to distinguish nest building from the main forms of maternal behavior: (*a*) Inorganic material was dropped in the nest and disregarded, whereas a young frequently elicited, besides retrieving, a variety of activities from a maternal mouse: it was sniffed at rapidly, the genitalia were smelled, and then, once in the nest, it was turned between the paws, thoroughly licked and cuddled. (*b*) As already stated, young presented to a maternal animal elicited retrieving, licking, and cuddling regularly, but rarely digging and carrying of inorganic material. (*c*) Wiesner and Sheard (34) observed rats that "retrieve young eagerly, but do not build a nest, while in others the reverse holds true." (*d*) A maternal behavior, with its typical discrimination between young of different ages and preference for the youngest ones, seemed to be present in most species, as revealed by casual observations in cats, dogs, and even in man. Usually, however, such a maternal behavior occurred independently or in the absence of nest building. (*e*) Furthermore, the intensity of retrieving and discrimination

between young of different ages was comparable in the lactating females and in parental males (Table 1), although the former built nests actively and the latter did not

One may conclude that a drive, specifically directed towards the young, determined the main forms of maternal behavior in the mouse

The action of hormones on such behavior is fairly certain in several species, for instance, in the dog. When a female dog in heat is not sired by a male, a pseudo-gestation follows the heat period (19) At the end of this pseudo-gestation, lactation (19) and maternal behavior (17) takes place, as after a normal gestation, although there is no parturition and no young sucking Such an occurrence of maternal behavior can only be explained by the hormonal shifts present at the onset of lactation

In the mouse the only features to suggest a hormonal influence on maternal behavior were the increase in the maternal activity of the females (*N* and *F* strains especially) before or at the time of the first parturition, and the slightly more intense maternal behavior in multiparous than in primiparous females. In general, during or after the lactation period, no variations of maternal comportment could be found The rôle of hormones is therefore very restricted in the mouse.

What hormones act on maternal behavior has not so far been conclusively demonstrated. Two theories based on the two main hormonal features of lactation, namely secretion of prolactin by the hypophysis and interruption (or reduction) of the ovarian-secretion of female hormone, estrin, have been presented. (*a*) Ehrhardt (7), Wiesner and Sheard (34), and McQueen-Williams (18) attributed maternal behavior to secretions from the hypophysis. The active factor in the hypophysis, after Riddle, Bates, and Lahr (26, 27) and Noble, Kumpf, and Billings (21) would be prolactin. (*b*) According to Ceni (3), maternal behavior is due to a reduction of the gonadal secretions. However, as the secretion of prolactin and the absence of estrin secretion are linked (19, 6, 1) and both occur during lactation, it is not possible to decide which one of these two conditions has a direct bearing on maternal behavior.

In the mouse, the presence of prolactin or the absence of estrin might be the effective hormonal agent, but neither was essential for maternal behavior. The suppression of prolactin, by removal of the

hypophysis, did not prevent the development of maternal activities. Nor did the injection of generous amounts of female hormone for periods of several months affect the occurrence of maternal behavior.

Riddle stated that several hormones, namely, estrin, progesterone and prolactin, must act successively in the body before maternal behavior appears (rat, chicken) (28). However, in mice which were castrated at birth to prevent the sex hormones from acting as a priming step, maternal behavior still occurred after sensitization.

Another possible intervention of hormones resulted from the fact that it was necessary to leave a young in the cage for some time before maternal behavior appeared in non-maternal animals. As happens often in birds, sensorial stimuli, which here would come from the young, might have determined the release of a hormone from the hypophysis of the mother. Thus in the pigeon the presence of the eggs in the nest conditioned the release of prolactin from the hypophysis of the parents, since the removal of the eggs prevented the prolactin-induced development of the crop gland (23, 25). In the mouse, the presence of young, being effective in animals deprived of their hypophyses, did not act through a release of a hypophyseal hormone.

Finally the presence of maternal behavior in males eliminated the possibility of the behavior being caused by the reflexes of parturition or of nipple suction. The latter, however, had an action in lactating females which will be referred to later.

As a consequence of the restricted action of the hormones in the maternal comportment of the mouse, a great importance must be given to the rôle of the nervous system. It is well known that the pattern of a behavior is usually dependent upon neuronal mechanisms (4). Hence the existence of nervous elements commanding maternal behavior could be assumed. The localization of these elements might be in the corpus striatum. Various lesions of this region in the pigeon (29) and in the cichlid fish (22) appeared to suppress one, several, or all the maternal activities of the animals. In addition, the whole cerebral cortex had an influence in mammals, mainly on the correlation of the units of behavior (2).

In the initiation of maternal activity observed at the first parturition of the female mouse, the active factor, in addition to a slight hormonal stimulation, may be a "sensitization" of the maternal behavior of the mother by her young. Experimentally, such "sensi-

tization" could be easily obtained by putting new-born young in the cage of animals. This sensitization, being similar in normal, castrated, or hypophysectomized animals, obviously involved only nervous mechanisms.

For similar reasons, the provocation of maternal behavior in animals which were already parental may be ascribed to nervous or sensorial mechanisms. Further proof of this fact was obtained from a comparison of the responses of lactating females and of parental males presented with young of various ages. The provocative values of the young were not greater when tested on females during lactation than when tested on sensitized males (Table 1). Therefore, the provocation of maternal behavior depended on nervous mechanisms apparently not influenced by sex or lactation.

Finally, maternal behavior once initiated was maintained for periods of time as long as three months even in the absence of young (Table 2). This maintenance also appeared to depend on the nervous system.

In conclusion, the nervous mechanisms that are responsible for the maternal comportment may function independently of hormones not only in parental animals, but also in non-parental animals, when the sensorial stimulus afforded by the young is prolonged for some time, as in the sensitization process.

3. *Other Components of Maternal Behavior*

a. Nest building was elicited by cold even in non-lactating animals (Kinder, 11). The action of cold was immediate, nest building being observed as soon as the animal was put in a colder place (16). According to Kinder, nest building is "a reflex of cold" as are shivering and goose flesh in man (11). During lactation, the cause of nest building might be the "reflexes of cold" due to the supernormal temperature found in lactating animals (16). Such reflexes are known in cases of fever in man, especially when the temperature is variable. That the high temperature of lactating rats was unstable supported the opinion that reflexes of cold occurred which determined nest building. Factors other than this lactation fever might play a rôle, since a high, variable temperature persisted all during lactation, while nest building gradually decreased (16).

b. The prolonged station on the nest. The attraction for the young which accompanied the main forms of maternal behavior ex-

plained why the mother went over the young in the nest. The persistent station on the nest for long periods of time was a consequence of the marked reduction in the activity of the lactating mother (32). On the other hand, the relaxed attitude of the mother in the act of nursing suggested to us the existence of a pleasure derived from suckling the young. The theoretical necessity of such a pleasure for the preservation of the species and its existence in other animal species (cat) and even in women pointed to the correctness of such an assumption. If present, a pleasure derived from suckling would contribute in keeping the mother over the young.

c. *The defense reactions*, being observed for the most part during the act of nursing, appeared as another consequence of a pleasure derived from nursing, since they compared with defense reactions present in animals satisfying other drives (e.g., dog when feeding).

Our account of the initiation and maintenance of the pattern of maternal behavior in lactating female mice may be briefly summarized: (*a*) Their intense *nest building* was, at least in part, determined by the reflexes of cold that occurred as a consequence of a hypernormal and unstable temperature. (*b*) Both reduction in the activity of the animals and pleasure derived from nursing explained the *prolonged station on the nest*. (*c*) *The defense reactions* often seem to accompany the intense phases of various drives. (*d*) Finally, *retrieving, licking,* and *cuddling* of the young, which were looked upon as the main components of maternal behavior, were due in part to the "sensitization" of the mother afforded by the continuous presence of the young after parturition. Evidence that some stimulation of maternal behavior by hormones existed at parturition was reported above, but the rôle of the hormones in the mouse was less important than in the rat and apparently most other species. A marked maternal care, being realized through sensorial or hormonal stimuli, will persist indefinitely without need of further stimulus.

E. Conclusions

The complete pattern of maternal behavior in rats and mice involves elements of diverse origins. Nest building, prolonged station on the nest, and defense reactions are caused by factors without direct relation to the main features of maternal behavior, namely retrieving, licking, and cuddling of the young.

1 In maternal mice, the latter behaviors are elicited by young and the ability of the young to do so decreases progressively as their age increases

2 The main expressions of maternal behavior in mice are little influenced by the hormones of parturition and lactation. The only features probably to be attributed to hormonal influences are an increase in maternal activities just before the first parturition and slightly greater activities in multiparous than in primiparous females No other variations could be detected either during or after lactation periods

3 In normal male and virgin female mice, even after removal of sex hormones by castration or hypophysectomy, maternal behavior could be evoked by leaving young in the cage (sensitization)

4. Nervous mechanisms may therefore be considered the primary factor involved in the initiation and maintenance of the main features of maternal behavior.

5. An explanation of the complete pattern of maternal behavior in post-partum females, on the basis of physiological findings, has been attempted.[5]

References

1 BATES, R W, RIDDLE, O, & LAHR, E L. The mechanism of the antigonad action of prolactin in adult pigeons *Amer J Physiol*, 1937, **119**, 610-614

2 BEACH, F A The neural basis of innate behavior. II Relative effects of partial decortication in adulthood and infancy upon the maternal behavior of the primiparous rat *J Genet Psychol*, 1938, **53**, 109-148

3 CENI, C Sulla transformazione dell istinto sessuale in istinto materno nella femina e nel maschio Ricerche sperimentali *Proc 2d Internat Cong Sex Res*, London, 1930

4 COGHILL, G E Anatomy and the Problem of Behavior Cambridge University Press, 1929

5 COX, W M, & MUELLER, A. J The composition of milk from stock rats and an apparatus for milking small laboratory animals. *J Nutrit*, 1937, **13**, 249-260

6 DRESEL, I The effect of prolactin on the estrous cycle of non-parous mice *Science*, 1935, **82**, 173

[5]The author wishes to thank Dr H. W. Nissen for suggestions and correction of the manuscript, and Doctors Edgar Allen and W O Nelson for advice and assistance.

7. Ehrhardt, K. Beitrag zur Hypophysen-Vorderlappenreaktion unter besonderer Berücksichtigung der Aschheim-Zondekschen Schwangerschaftsreaktion. *Klin. Woch.*, 1929, **8**, 2044-2046.

8. Ellis, H. Studies on the Psychology of Sex. IV. Sexual Selection in Man. Philadelphia: Davis, 1927.

9. Giard, A. Les origines de l'amour maternel. Oeuvres diverses, Vol. 1, 215.

10. Inglebrecht, P. Influence du système nerveux central sur la mamelle lactante chez le rat blanc. *C. R. Soc. Biol.*, 1935, **120**, 1369-1371.

11. Kinder, E. F. A study of the nest-building activity of the albino rat. *J. Exper. Zool.*, 1927, **47**, 117-161.

12. Lahr, E., & Riddle, O. Temporary suppression of estrous cycles in the rat by prolactin. *Proc. Soc. Exper. Biol. Med.*, 1936, **34**, 880-883.

13. Leblond, C. P., & Nelson, W. O. Maternal behavior in hypophysectomized male and female mice. *Amer. J. Physiol.*, 1937, **120**, 167-172.

14. Leblond, C. P. Extra-hormonal factors in maternal behavior. *Proc. Soc. Exper. Biol. Med.*, 1938, **38**, 66-70.

15. ——— L'instinct maternel: Nature et relations avec la glande mammaire, l'hypophyse et le système nerveux. *Rev. franç. d'Endoc.*, 1937, **15**, 457-475.

16. ——— Unpublished results.

17. Loisel, G. Relations entre les phenomènes du rut, de la lactation, de la mue et de l'amour maternel chez une chienne hybride. *C. R. Soc. Biol.*, 1906, **60**, 255-258.

18. McQueen-Williams, M. Maternal behavior in male rats. *Science*, 1935, **82**, 67-68.

19. Nelson, W. O. Endocrine control of the mammary gland. *Physiol. Rev.*, 1936, **16**, 488-526.

20. Nissen, H. W. A study of maternal behavior in the white rat by means of the obstruction method. *J. Genet. Psychol.*, 1930, **37**, 377-393.

21. Noble, G. K., Kumpf, K. F., & Billings, V. N. The induction of brooding behavior in the jewel fish. *Endocrinology*, 1938, **23**, 353-359.

22. Noble, G. K. The function of the corpus striatum in the social behavior of fishes. *Anat. Rec.*, 1936, **64** (Supp.), 34.

23. Patel, M. D. The physiology of the formation of the pigeon's milk. *Physiol. Zool.*, 1936, **9**, 129-147.

24. Richter, C. P., & Barelare, B. Nutritional requirements of pregnant and lactating rats studied by the self-selection method. *Endocrinology*, 1938, **23**, 15-24.

25. Richter, C. P. Hypophyseal control of behavior. *Symposia on Quantitative Biology*, Cold Spring Harbor, July, 1937, 258-268.

26. Riddle, O., Bates, R. W., & Lahr, E. L. Prolactin induces broodiness in fowls. *Amer. J. Physiol.*, 1935, **111**, 353-358.

27. Riddle, O., Lahr, E. L., & Bates, R. W. Maternal behavior induced in virgin rats by prolactin. *Proc. Soc. Exper. Biol. Med.*, 1935, **32**, 730-734.

28. Riddle, O. Aspects and implications of the hormonal control of the maternal instinct. *Proc. Amer. Philos. Soc.*, 1935, **75**, 521-525.

29 ROGERS, F. T Studies of the brain stem: An experimental study of the corpus striatum of the pigeon as related to various instinctive types of behavior *J. Comp Neurol*, 1922, **35**, 21-51.

30. SELYE, H., COLLIP, J B, & THOMSON, D L Nervous and hormonal factors in lactation *Endocrinology*, 1934, **18**, 237-248

31. SELYE, H, & McKEOWN, T Further studies on the influence of suckling *Anat Rec*, 1934, **60**, 323-332.

32 SLONAKER, J R The effect of copulation, pregnancy, pseudo-pregnancy, and lactation on the voluntary activity and food consumption of the albino rat *Amer J Physiol*, 1925, **71**, 362

33 STURMAN-HULBE, M, & STONE, C P Maternal behavior in the albino rat. *J. Comp Psychol*, 1929, **9**, 203-227

34. WIESNER, B P, & SHEARD, N M Maternal Behavior in the Rat London Oliver & Boyd, 1933

35. WITSCHI, E Seasonal sex characters in birds and their hormonal control *Wilson Bull*, 1935, **47**, 177-188

Laboratoire de Synthèse Atomique
67 Rue Franklin
Ivry (Seine), France

The Journal of Genetic Psychology, 1940, 57, 345-391

THE RESULTS OF MULTIPLE BINET RETESTINGS OF THE SAME SUBJECTS THE EDUCATIONAL IMPLICATIONS OF VARIATION OF TEST PERFORMANCE*[1]

Department of Special Education and Mental Hygiene, Wilmington, Deleware

J. E. W. WALLIN

A. INTRODUCTION

Prior to 1909 psychological examination techniques played a negligible rôle in the selection of children for the special classes for the mentally retarded which had been established theretofore in the public schools in the twenty odd cities in this country which supported such classes at that time.[2] In most instances the selections were based primarily on the child's scholastic record, sometimes supplemented by the results of medical examinations. A few home-spun, unstandardized, common sense tests and possibly a few of the simpler tests familiar to the psychological laboratories were, no doubt, toyed with in the two public school child study departments then in existence, the one in Chicago, established in 1898, and the one in Rochester, established in 1907, and in the few existing extra-school clinics which attempted to cooperate with public school authorities in pupil classification. But the sum total of such sporadic and largely amateurish testing, whatever its value, available to the schools that maintained special classes was negligible before the

*Received in the Editorial Office on September 8, 1939.

[1]Presented, in part, before the Detroit meeting of the International Council for Exceptional Children, February 24, 1939.

[2]The pioneer special classes for mentally deficient or retarded children were established in the following cities in the order mentioned. Providence, 1896, Springfield, Mass., 1897, Chicago, 1898, Boston, 1899, New York, 1900, Philadelphia, 1901; Los Angles, 1902, Detroit and Elgin, 1903; Trenton, 1905, Washington, D. C., Bridgeport, Conn., Newton, Mass., and Rochester, 1906, St. Louis, Indianapolis, Passaic, New Britain, and Saginaw, 1907, Milwaukee, Reading, and Houston, 1908. Additional historical data on the early special classes of various kinds and on the pioneer psychological and psychoeducational clinics and child study departments in higher institutions of learning, public schools, residential institutions, and courts are available in the writer's *The Mental Health of the School Child*, pp. 22-114, and 383-428, 1914.

appearance of Goddard's translations of the De Sanctis and the Binet-Simon tests of general intelligence in December, 1908, and the subsequent Vineland revision of the Binet scale in 1911. These inconspicuous events, perhaps known at the time to only a few in this audience, marked a definite turning point in the history of the public-school special-class movement in the United States. After the appearance of the Vineland publications, especially the 1911 scale, intelligence testing spread like wild fire in this country in the public schools, courts, residential institutions for dependents and delinquents, and elsewhere. Within a few years 30,000 copies of the Vineland Binet manuals had been distributed. Although the enormous spread of the scale was partly due to the exaggerations promulgated regarding its unerring accuracy and its complete adequacy for the diagnosis of mental defectiveness (we were assured that "nothing else was required in the large majority of cases"), it was also due to the fact that the scale met a genuine need better than had ever been done before, namely, the need for a simple, practical, easily administered, standardized, objective scale of tests for the measurement in understandable terms of levels of intelligence and degrees of intellectual deviation. Indeed it was the need in the Parisian schools for such an instrument that led to its invention. It is in large measure due to the Binet scale that psychological examinations prior to assignment to special classes has become a commonplace practice throughout the nation. Perhaps it is not an overstatement to affirm that all of our modern psychological testing in the schools for purposes of educational classification or mental diagnosis originated in the Binet scale, for group mental testing did not flourish on any scale until 1918 or 1919.

In spite of the numerous criticisms lodged against the Binet scale during the three decades in which its various revisions have been in constant use—some of which have been superficial or irrelevant while others have been wholly pertinent and in some instances even devastating in their implications—it probably is correct to say that the Binet scale has been used more frequently than all other clinical tests combined for the determination of the level of general intelligence, certainly so far as concerns school children, and that it still continues to occupy this position of preeminence. The continued popularity of the scale, whatever its theoretical shortcomings and practical imperfections, is attested by the fact that two American

revisions have appeared within the last two years, neither of which marks any basic reconstruction.

Let no one infer from this brief introduction that we are preparing the ground for the presentation of evidence that the Binet scale is the final word in intelligence measurement technique, that the new revisions cannot be improved, that the system of tests or of scale construction is based on correct psychological principles, or that the numerous questions brought to the fore by the use of the scale are now fully understood and have been satisfactorily solved, such as the questions of the relative influence of hereditary and environmental factors on the test results the nature, extent, and causes of fluctuations in successive test performances, the nature and causes of irregularities or scattering in the pattern of response to the individual tests, the reliability of the test scores for diagnosis or prognosis, their educational and social implications, and the like. On the contrary, few of the numerous questions propounded by the prodigious Binet measurement work of more than a quarter century have been settled with finality, although during this interval more critical studies have probably appeared of the different Binet revisions and of results obtained from their use than of all other individual tests currently used in clinical practice. The most recent criticisms by the Iowa Child Welfare Research Station workers are as challenging as any that have been leveled against the scale during its entire history, although the assault on the scale itself is implicit rather than direct and perhaps less affects its fundamental structure than the uses made of its results as well as of the results of other tests of ability These investigators' "pessimistic commentary on the value of intelligence tests" seems to lead inescapably to the conclusion that a Binet *IQ* affords a reliable measure of intelligence only at the time the test is given and possesses no prognostic value as a measure of inherent or constitutional endowment It affords a measure of environmental stimulation rather than of native ability To quote "The *IQ* of a child at a given age does not necessarily tell us what his future *IQ* will be or what his past *IQ* was" "There is now no escape from the fact that the *IQ*'s of children have possibilities of change over a large portion of the *IQ* range from genius to feeblemindedness. Children can and do change in test *IQ* from average to genius and from average to feeblemindedness," depending upon the stimulating or non-stimulat-

ing character of the environment. Instances are cited by the investigators of Binet *IQ*'s which increased from 98 to 167, an increase of 69 *IQ* allegedly due to "attendance at preschool and enrollment in an unusually stimulating elementary school," and which decreased from 103 to 60, a loss of 43 points allegedly due to an "unfavorable environment . . in an orphanage where almost no individual attention was given the children, where as many as 35 children were housed in one cottage with an untrained matron in charge, where play materials were not available, where contact with ordinary life situations was at a minimum" (7, 8, 2, 3) These startling conclusions regarding the possibilities of education may give pause to many special educationists and clinical psychologists, particularly those who have excluded entirely from the benefits of public education children adjudged ineducable because they fall below an *IQ* of 50, which they regard as indubitable proof of an ineradicable or unalterable native defect On the other hand, the conclusions may arouse the skepticism, not to say the undisguised hostility, of the host of workers who have striven to provide mentally retarded children with the best possible educational facilities that modern school systems afford but who have never been able, figuratively speaking, to transform sow's ears into silken purses and who are unable to follow the conclusions repeatedly promulgated by practically all of the Iowa group of investigators.

For a long time we have given lip service to the concept of feeble-mindedness from social isolation but, practically, the concept has remained a dead letter. The majority of specialists on mental defect do not even mention the possibility of mental defectiveness through social deprivation Apparently the assumption has been accepted that environments so lacking in intellectual stimulation as to produce feeble-mindedness do not exist on the reality plane in this country or that feeble-mindedness engendered by a socially impoverished environment is a myth. Of many thousands of cases passing through the writer's clinics only four instances of serious mental retardation now come to mind which were ascribed purely to mental or educational neglect The case files for these adolescents are not now at hand, but the *IQ*'s based on the 1911 Vineland Binet varied from about 60 to 72. Two unschooled sisters, who had been put to work in a mill in Florida at the age of six, were brought by a religious organization to Pittsburgh at the ages of 10 and 14

for an education, and placed in a child welfare home. While testing mentally deficient, they did not present quite the symptom complex of the mental defective. The same can be said of a 19-year-old boy from the same state who had never been in school, who reached my clinic in St. Louis, referred from a first grade, at the age of 18 when he had a Binet level of about nine and was diagnosed as a case of mental defect through cerebral atrophy superinduced by social and educational neglect, and of a 14- or 15-year-old girl, also without any schooling, from a backwoods mountain section in West Virginia. None of these would rate below the high grade defective or borderline level. Although few authorities have made any use of this concept, we are now told by the Iowa investigators that whether children of the same initial *IQ* are destined for feeble-mindedness or for genius depends almost solely upon the character of the environmental stimulation to which they are subjected. I can see in my fancy many an incredulous Thomas shaking his head and protesting vociferously that he cannot accept the conclusion that the environment is omnipotent in the development of intelligence and that heredity counts for naught. Let us hope that such a doubter will some day repeat the Iowa experiment for purposes of verification and that the results will turn out differently from the results of the British investigators who repeated Stockard's experiments on the parental alcoholization of guinea pigs and were unable to substantiate his cacogenic findings. Certainly the Iowa findings if verifiable are of transcendent importance for education so far as their implications are concerned. But all the remarks thus far made are merely preliminary to the main contribution of this paper.

B. Description of the Investigation

The investigation now to be reported is unique, it is believed, in at least two particulars. First, it is based on the largest number of Binet tests periodically administered to the same subjects, so far as is revealed by the published results of such testings that have come to the writer's attention. Twenty-seven semi-annual Stanford-Binet (1916 version) testings were administered to Subject *A* and 26 to her sister, Subject *B*, four years and three months younger than *A*. If the six easiest tests from the 1905 Binet scale and the easiest tests in the Stanford-Binet, which were administered from the 12th to the 30th month, are included, the number of testings

is 31 for Subject *A*, and 29 for *B*. The data derived from such an extensive, unduplicable series of tests of two normal siblings should possess intrinsic interest, whether or not the Binet scale has been outmoded, as some critics aver, and quite apart from any bearing the results may have on many of the moot issues affecting intelligence testing.

Second, all the testings were done by the same person. This should make for uniformity of test administration and scoring. Binet tests given to the same child by different examiners show a correlation of .79 as compared with .87 when given by the same examiners (Gertrude Hildreth).

In the administration of the Binet scale, the wide range method of testing was employed. Thus, to illustrate from two age levels: at age five, *A* was given 21 tests above her V-year base to age IX (exclusive of the vocabulary test in X counted at VIII) and *B* was given 24 tests above her IV-year base to age IX (exclusive of the vocabulary test in X). At age 9, both *A* and *B* were given 20 tests above the base of VIII (counting the vocabulary tests in both X and XII). This does not include the tests given at and below the base. Many tests were administered many times during the course of the investigation. The extent of the testing during any sitting was determined by the nature of the performance in the upper year-levels.

In addition to the Binet, the examinations included the more or less frequent administration of the Seguin form board, given as a test of maximum speed of reaction, and Healy *A* and Healy *B* form boards, the Wallin peg boards, and certain anthropometric measurements (height, weight, spirometry, manometry, and cephalometry). In addition to the tests, a large number of observations were recorded, mostly by the mother, particularly during the first few years of life, of many behavior characteristics, the development of crawling, sitting, standing, walking, and speech, dentition, illnesses, accidents, falls, and the like. These observations were not made under highly controlled laboratory conditions, in consequence of which some of them are more or less fortuitous and unsystematic. Nevertheless, they possess a certain counterbalancing advantage in that they record happenings in relatively free and unregimented life situations in a presumably normal home environment.

The time available for this paper limits the presentation to an

TABLE 1

STANFORD-BINET AGES AND BINET INTELLIGENCE QUOTIENTS FOR EACH TEST ADMINISTRATION

Chron age	Subject *A* Binet age	Subject *A* *IQ*	Subject *B* Binet age	Subject *B* *IQ*
3[1]	3- 2	106	3- 6	117
3½	4- 0	114	4- 1	124
4	4- 4	108	4-10	121
4½	5- 2	115	5- 4	116
5	6- 0	120	5-10	117
5½	6- 6	118	6- 4	113
6	7- 2	119	6-10	114
6½	7- 4	113	7- 1	109
7	8- 2	115	8- 6	121
7½	8- 4	111	9- 1	120
8	9- 4	117	9- 4	117
8½	9- 2	108	10- 6	122
9	9-11	110	10- 8	119
9½	10- 9	112		
10	11- 4	112	11- 4	113
10½	11-10	113	11- 7	110
11	12-11	117	11- 7	105
11½	12- 6	108	12-11	112
12	14- 2	118	13- 4	111
12½	15- 5	123		
13	15- 8	121	15- 4	118
13½	15- 9	117		
14	16- 7	118	15- 4	110
14½	17- 6	125[2]	16- 8	119[2]
15	17- 2	122[2]	16- 9	120[2]
15½	18- 6[3]	133[2]	16- 8	119[2]
16	18- 0[2]	129[2]	17- 1	122[2]
16½			17- 6	125[2]
17			18- 6[3]	132[2]
14½	17- 6	121[4]	16- 8	115[4]
15	17- 2	114[4]	16- 9	112[4]
15½	18- 6[3]	120[4]	16- 8	108[4]
16	18- 0	113[4]	17- 1	107[4]
16½			17- 6	106[4]
17			18- 6[3]	116[4]

[1]The tests sometimes could not be given on the exact birthdays or half-birthdays because of various interferences The average deviation from the exact birthdays was 7 6 days for *A*, irrespective of direction (all plus deviations except four) and 11 8 for *B* (all plus)

[2]Age 14 used as the divisor

[3]By failing one test in age XVIII in the Stanford-Binet Scale (the forward digits test was failed by *A* at 15½ but passed at 16 while *B* failed on this test throughout and after passing all the tests in XVIII) it is possible to get a six months' higher rating than by passing all the tests, as each test has a value of 6 months in age XVIII *A* base of XVI and five

passes (30 months) in age XVIII yield a rating of 18-6. Such records doubtless yield spurious IQ increases.

[4]Based on the actual chronological age except above 16-0, when 16 is used as the divisor.

Before the age of three the six easiest tests in the Binet 1905 scale and the tests in the Stanford-Binet ages III, IV and V were administered, with the following results expressed in terms of the number of tests passed (rather than in terms of Binet ages):

Age in months	Subject A	Subject B
12½		5 (including one questionable response)
13	5	
18	6	
24	8	9 (including one questionable response)
30	12 (including 2 questionable responses)	11 (including one questionable response)

untechnical recital of some of the more interesting or important findings and of some of the major implications. Because of the necessary brevity, rather complete tabulations have been supplied for the benefit of those who may be interested to analyze the data more fully than could be done here. The tables include the Binet age and IQ for each test (Table 1), the Seguin scores (Table 2) and peg board scores (Table 3) for each test, the mean Binet IQ's for the whole series of annual and semi-annual testings and for various age periods (Tables 4 and 5), the means for the Seguin form board for two age periods (Table 6), the SD's, MV's, Q's, PE's and C's for the whole series of annual and semi-annual testings and for various age levels (Table 7), various developmental data (Table 8), the record of illnesses and injuries (Table 9), and the school marks (Table 10).

C. Results

Many striking similarities and certain striking differences stand out in the records of these two siblings who possess few of the mental or physical traits characteristic of identical twins or identical siblings. The older one is spare while the younger one is stout. The older tends toward introversion, the younger toward extroversion. The older tends to be somewhat inhibited and emotionally restrained while the other tends to be somewhat effusive emotionally. The older is not inclined to be a free mixer while the younger one readily establishes social contacts.

TABLE 2

SEGUIN FORM BOARD RECORDS FOR EACH TRIAL

Age in years	Subject *A* Fastest[1]	Ave[2]	Subject *B* Fastest	Ave
2	113[3]			
2½	41.2	58.6	Failed[4]	
3	40.4	50.3	65.0	99.3
3½	32.6	36.1	50.0	76.0
4			30.0	42.3
4½	23.0	27.9	28.2	33.2
5			22.2	25.8
5½	21.0	38.6		
6			15.4	19.1
6½	20.8	21.4	18.0	19.4
7	18.6	21.6	20.0	24.0
7½	17.4	18.8	17.4	19.1
8	20.0	20.0	12.5	15.8
8½	14.0	16.0	18.4	21.0
9½	12.4	17.4		
10	12.0	15.1	11.8	15.4
10½	14.0	15.1		
11	13.0	13.8	12.0	12.1
11½	13.4	15.1		
12½	14.2	15.1		
14	11.6	15.5		
14½	12.0	12.7		

[1]Fastest time in seconds of three trials

[2]Ave, average time for three trials, except for *A*, age 8 (two trials)

[3]Failed on the first and second attempts

[4]Random hit and miss attempts, failed after 172 seconds in the first trial and after 196 seconds in the second attempt

1. *Similarities in Test Findings*

Surprisingly, the average *IQ* for the entire series of semi-annual Stanford-Binet testings based on age 14 as the highest divisor for ages above 14 is 116.3 for *A* and 117.1 for *B*, a negligible superiority of 0.8 *IQ* for *B*. This is based on the average of all the tests from ages three to 16 for *A*, the senior sibling, who at 16 had successfully performed all the tests, except the XVIII-year forward digits test, and from ages three to 17 for *B*, who at 17 had completed all the tests except the forward digits test. On the other hand, on the basis of the use of age 16 as the highest divisor for ages above that base, *A*'s average *IQ* is 114.8 and *B*'s 114.3, a negligible superiority of 0.5 *IQ* for *A*. Because of the varying practice among examiners in the use of the two divisors, 14 and 16, for children of 14 and

TABLE 3
WALLIN PEG BOARD

Age in months	Subject A Board A F[1]	Subject A Board A Ave	Subject A Board B F	Subject A Board B Ave	Subject A Board C F	Subject A Board C Ave	Subject A Board D F	Subject A Board D Ave	Subject B Board A F	Subject B Board A Ave	Subject B Board B F	Subject B Board B Ave	Subject B Board C F	Subject B Board C Ave	Subject B Board D F	Subject B Board D Ave
12½									Played with the pegs for about 10 m Kept removing pegs, allowing only one to remain in place at a time, except once when 4 were in place							
18	32	39.3	Failed after 3 m twice[2]		Failed after 7 m twice											
21	19.5	28.2[3]	21.0	50.0[4]	27.0	34.6[5]	27.0[6]									
24									21.2	23.4	35.6	39.5	77.0[7]		131.0[8]	
30									11.0	13.3	18.4	28.1	31.4[9]		Failed[10]	
36									18.0	19.8	18.0	18.3	18.0	37.3[11]	27.0	55.8[12]
42									14.5	14.7[13]	13.4	21.7[14]	19.0	22.3	21.0	21.2[15]
48									11.0	12.5	15.4	15.6[12]	12.4	12.9	18.0	19.5[16]
54	10.2	11.8	12.2	12.5	11.0	12.6	21.8	28.0[16]	11.6	12.4	11.2	14.7	11.0	13.6	13.2	14.4

[1]Fastest reaction in two or three successes

[2]Inserted the pegs correctly but did not permit them to remain.

[3]Played with the pegs the second trial, distracted by noise the third time

[4]Crude coördination.

[5]Interested, attentive

[6]Failed second and third trials—distracted by children, failed on triangles third trial

[7]One success after correcting two errors

[8]One success

[9]Failed second and third trials

[10]Largely hit and miss responses, but had five correct second trial

[11]Failed the second time, confused squares and cylinders but not triangles

[12]Delayed second trial because of removal and replacement of pegs

[13]Based on 2 trials

[14]Two trials, distracted the second time

[15]Two trials

[16]Lack of interest

TABLE 4

AVERAGE INTELLIGENCE QUOTIENTS FOR THE ENTIRE RANGE OF STANFORD-BINET TESTING AND FOR ARBITRARILY DETERMINED SUCCESSIVE AGE PERIODS FROM 3 TO 7, 8 TO 12, AND 13 TO 16 OR 17

	Subject A	Subject B
Annual testing, ages 3 to 16 or 17[1]		
Based on age 14 as highest divisor	116.5	117.1
Based on age 16 as highest divisor	114.8	114.5
Semi-annual testing, ages 3 to 16 or 17[1]		
Based on age 14 as highest divisor	116.3[2]	117.1[3]
Based on age 16 as highest divisor	114.8[4]	114.3[5]
Ages 3 to 7	113.9	118.2
Ages 8 to 12	113.8	113.6
Ages 13 to 16 or 17 based on age 14 as highest divisor	123.4	120.6
Ages 13 to 16 or 17 based on age 16 as highest divisor	117.7	111.5

[1]Age 17 in the case of Subject B
[2]Median 116
[3]Median 117.5
[4]Median 114.5
[5]Median 114.5

TABLE 5

AVERAGE BINET-SIMON INTELLIGENCE QUOTIENTS FOR VARIOUS AGE PERIODS IN WHICH ONE OR THE OTHER SUBJECT EXCELS

Age period	Subject A	Subject B
From 3 to 4½	110.7	119.5
From 5 to 6½	117.2	113.2
From 7 to 10	112.1	118.6
From 10½ to 16 or 17, based on age 14 as divisor	120.1[1]	114.6
From 10½ to 16 or 17, based on age 16 as divisor	115.9	110.8

[1]The Binet age at 15½ is counted as 18-0 instead of 18-6 because of the spurious increase as explained in Table 1. The averages in this table are based on the ages in which the tests were administered to both subjects.

TABLE 6

AVERAGES FOR THE SEGUIN FORM BOARD FOR SUCCESSIVE LIFE PERIODS

	Subject A		Subject B	
Interval	Fastest trial	Ave. of three trials	Fastest trial	Ave. of Three trials
3 to 4½	32.0	38.1	43.3	62.7
7 to 10	15.73	18.15	16.02	19.06

TABLE 7
VARIABILITY BETWEEN SUCCESSIVE BINET IQ'S

Series	Subject A SD[1]	C[2]	PE[3]	SD[4]	MV[5]	Q[6]	Subject B SD[1]	C[2]	PE[3]	SD[4]	MV[5]	Q[6]
Annual testing, ages 3 to 16 or 17[7]												
Based on age 14 as highest divisor	6 06	052	4 08	5 83	4 57		6 26	053	4 22	6 05	4 40	
Based on age 16 as highest divisor	4 57	039	3 08	4 40	3 74		4 80	041	3 23	4 65	3 96	
Semi-annual testing, ages 3 to 16 or 17[7]												
Based on age 14 as highest divisor	6 39	054	4 31	6 27	4 98	4 2	5 86	050	3 95	5 74	4 53	4 0
Based on age 16 as highest divisor	4 56	039	3 07	4 47	3 78	3 2	5 20	045	3 50	5 11	4 34	3 7
Ages 3 to 7	4.58	040	3 08	4 34	3 72		4 46	038	3 00	4 23	3 44	
Ages 8 to 12	4 84	042	3 26	4 60	3 96		5 44	.047	3 67	5 09	4 27	
Ages 13 to 17[7] based on age 14 as highest divisor	5 56	045	3 75	5 15	4 48		6 27	051	4 22	5 87	4 27	
Ages 13 to 17[7] based on age 16 as highest divisor	3 25	027	2 19	3 01	2 61		4 47	040	3 01	4 18	3 75	

[1]SD or σ (standard deviation) $= \sqrt{\frac{\Sigma(d)^2}{n-1}}$

[2]C (coefficient of variability) $= \frac{SD}{M}$ (based on the formula $\sqrt{\frac{\Sigma(d)^2}{n-1}}$

[3]PE (probable error) $= 0.6745\ SD$ (based on the formula $\sqrt{\frac{\Sigma(d)^2}{n-1}}$

[4]$SD = \sqrt{\frac{\Sigma(d)^2}{n-1}}$

[5]MV or AD (mean or average deviation) $= \frac{\Sigma D}{N}$

[6]Q (quartile deviation or semi-interquartile range) $= \frac{Q_3 - Q_1}{2}$

[7]Ages 3 to 16 in the case of Subject A and 3 to 17 in the case of B

TABLE 8

DEVELOPMENTAL DATA

	Age of crawling, creeping, or rolling
6 m 2 d	*A* rolled from one side of the bed to the other
7 m 22 d	*A* rolls from place to place on the floor
8 m 1 d	*B* made slight creeping movements
8 m 24 d	*B* crawls rapidly on knees and hands
10 m 11 d	*A* crawled on the floor for the first time
	Onset of walking movements and walking and running
3 m 7 d	*A* made walking movements when stood up—vigorous propulsive movements
6 m 27 d	*B* moves feet as if walking when stood up
9 m 28 d	*B* walked, holding to chairs.
11 m 3 d	*B* walked about five feet unsupported several times
11 m 6 d	*A* took one or two steps between chairs and mother several times
11 m 23 d	and 24 d. *B* took several steps unsupported
12 m 7 d	*A* took several steps unaided, running to mother
12 m 8 d	*B* ran several times the last few days
12 m 17 d	*A* walked unsupported three or four steps as never before
12 m 27 d	*A* ran long stretches without support
16 m 20 d	*B* stepped down the stairway for the first time
	Age of sitting
3 m 1 d	*A* tried to raise herself to sitting position, has sat up with pillow at back for several days
3 m 7 d	*B* almost raised herself to a sitting posture from a partially propped-up position
5 m 23 d	*B* sat a short time before falling backward
6 m 17 d	*A* sat alone on the floor for a minute
7 m 24 d	*A* sits on the floor unsupported for a while, usually eventually falling to the side
8 m 1 d	*B* sits with infrequent falls
8 m 10 d	*A* sits upright for longer intervals without falling
8 m 15 d	*B* raises herself to a sitting position from prone or supine position
9 m 16 d	*A* during the last few days has raised herself in bed to the sitting position
	Age of standing
1 m 27 d	*B* tries to raise herself when held in the hands.
4 m	*B* when held stood in the lap several minutes without bending knees.
4 m 25 d	*A* rises on feet when held in the lap, making vigorous reactions
5 m 24 d	*A* stood with feet separated
6 m 17 d	*A* stands on floor holding to mother
7 m 10 d	*B* stands on floor holding to something
10 m 26 d	*B* momentarily stood unsupported twice (and again at 11 m 5 d)
12 m 17 d	*A* stood unsupported with flour bag in hand
	Getting up and climbing up
12 m 24 d	*B* gets up from floor unassisted by using hands
13 m 9 d	*A* gets up from floor unaided by rising on hands
15 m 16 d	*B* got up on a small stool and then pulled herself up on the settee

TABLE 8 (*continued*)

16 m 28 d	*A* got up in the Morris chair and the rocker from the small stool
17 m 3 d	*B* pulled herself up on the swivel chair 16 inches high, stood up, leaned over the desk, and grabbed a pencil
18 m 7 d	*A* climbed up the big rocker alone
	Fingering and grasping
3 m 1 d	*A* fingered book and bath tub faucet
3 m 7 d	*A* made rudimentary attempts at hitting a rattle and grasping it
3 m 27 d	*A* began to grasp objects close by, reached for pocket book and tried to convey it to the mouth
4 m 5 d	*A* played with the bell on the rattle
4 m 5 d	*A* with both hands held the nursing bottle in the mouth
4 m 9 d	*B* has shown signs of grasping for several days, plays with rattle on the high chair
4 m 11 d	*B* plays with rattle in high chair
4 m 24 d	*B* holds the nursing bottle
	Smiling
24 d.	*B*, mother twice observed a facial reaction that resembled a spontaneous smile
44 d	*A* appeared to smile back at mother twice (these two facial reactions probably should be regarded as reflexes)
	Laughing
1 m 28 d	*A* laughed rather distinctly
2 m 1 d	*A* laughed aloud
2 m 21 d	*A* chuckled
3 m 3 d	*B* cooed and chuckled when talked to and smiled at
	Speech
	Pre-speech chatter or babbling
1 m 14 d.	*B* tried to talk
2 m 13 d	*A* tried to talk
	Reproduction of speech sounds
7 m 20 d	*B* repeated da da.
8 m 20 d	*A* repeated da da, ma ma
10 m 8 d.	*B* repeated da da, "ot dog," but not mama or papa
12 m 4 d	*A* repeated boo, ma, skat
	Meaningful expressions and short phrases
11 m 21 d	*B* said "howdy do" and "bye-bye"
11 m 23 d	*B* said "see that, see"
12 m	*B* said "see that," pointing Said "good-bye" when father left
13 m 1 d	*A* said "see"
15 m 9 d	*A* said "see the bear"
15 m 10 d	*B* said "good-bye" when father left
15 m 22 d	*A* said "see the boo woo"
17 m 10 d	*A* said "howdy do"
20 m	*B* said "open the door" [or gate] and "shut the door"
	Imitation of sounds of animals or of things
10 m 17 d	*B*, "What does the doggie say?" "Boo-woo,"
12 m. 4 d	*A* imitated "woo woo"
12 m 25 d	*B*, "What does the watch say?" "Tit, tit"
13 m 5 d	*A* can tell what the dog and the cow say Can now imitate a duck, a sheep, and a dog

TABLE 8 (*continued*)

	Evidences of alertness, intelligence, giving attention
1 m 18 d	*B* shows interest in new sights on railway journey
1 m 29 d	*A* fixedly looked at father's shiny shoes and also face at a distance of five feet
2 m 1 d	*B* has noticed her hands for several days
3 m 22 d	*A* was observed for the first time to be looking at her hands
2 m 18 d	*B*'s eyes followed father as he moved to the door
2 m 24 d	*A*'s eyes followed mother's movements
3 m 4 d	*B* for several days has followed her mother's movements by movements not only of her eyes but also of her head
4 m 5 d	*A* turns her eyes and body to follow father's retreating form
2 m 11 d	*B* recognizes mother's voice, turns head to find her when she speaks
2 m 23 d	*A*'s crying lessened when spoken to by father from an unseen position, turned head to locate the speaker
3 m 3 d	*B* seems to give more attention to sounds
5 m. 15 d	*A* gives more heed to voice sounds and whistling (also noticed at once lace on mother's new dress and tried to get hands on it, seems decidedly more alert)
25 m 16 d	*B* opened the door for the first time by turning the knob
33 m	*A* for the last two days has been able to open the door
	Dentition
6 m 14 d	*A*, two teeth erupted (lower central incisors)
9 m 12 d	*B*, lower central incisor erupted
7 m 8 d	*A*, upper left incisor erupted
12 m 8 d	*B*, upper left incisor appeared
7 m 24 d	*A*, upper right (lateral?) incisor out
11 m 5 d	*B*, upper right incisor.
14 m 21 d	*A*, two upper canines
16 m 18 d	*B*, upper canines

M, month, d, day

TABLE 9

RECORD OF COLDS, CHILDHOOD INFECTIONS, ILLNESSES, INJURIES, AND FALLS

Subject *A*		Subject *B*	
	Illnesses		
		6 wks	Severe cold
3 m 22 d	Cold	3 m 18 d	Severe cold for several days
3 m 24 d	Bad cold	4 m 14 d	Still coughs badly Fever for several days
4 m 19 d	Sneezing and coughing for a couple of days	4 m 17 d	Coughing spells Slight whooping cough
		4 m 21 d	Spots on tongue
6 m 1 d	Coryza during past week		

TABLE 9 (*continued*)

Subject *A*		Subject *B*	
	Illnesses (*continued*)		
6 m 9 d.	"Dreadful cold" last week		
6 m. 21 d	Sick most of week Green stools		
7 m 5 d.	Still not well. Lost 5 ounces last week Rather pale		
		9 m 12 d	Apparently had fever
11 m. 9 d	Fever	10 m 17 d.	Much coughing
13 m. 9 d	Severe cold	13 m. 28 d	Bad cold with cough Terribly frightened sometime ago by fire engine
		14 m 1 d	Given whooping cough antitoxin *A* had whooping cough Been coughing, not whooping
		14 m 16 d	Had fever a couple of days Coughing Tonsillitis
		14 m 20 d	Coughing some.
		15 m. 10 d.	Still coughs a little Whooping cough sign removed
		15 m 16 d	Little coughing now
		16 m 0 d	Seems to have become infected
17 m 10 d.	Little appetite for two weeks.	16 m 21 d	Bad cold a couple of days ago High temperature yesterday
19 m 21 d.	"Terrible cold"	18 m 28 d	Sick for two days
19 m. 24 d.	Quite sick Lagrippe		
19 m. 26 d.	Some better		
		24 m 14 d	Has had two bad colds
26 m 25 d	Fever today		
28 m. 17 d	Out on cold day without hat or coat	28 m 18 d	Earache for several days. High temperature
28 m 21 d.	Very sick, high fever, cough. Earache.	28 m 22 d	Eardrum punctured High temperature
32 m 10 d.	Given whooping cough serum.	30 m 2 d	Measles. Been incubating for some time Not as sick as *A* Good appetite after attack.
32 m 14 d	Vaccinated (did not take), given whooping cough serum		

TABLE 9 (*continued*)

	Subject *A*		Subject *B*
	Illnesses (continued)		
34 m 9 d	Much frightened in surf		
40 m 27 d	Cold Ear trouble	41 m 16 d	Treated for pyelitis
43 m 8 d	Eardrums contracted		
43 m 18 d	Spots on tonsils		
		54 m 11 d	Weak attack of chicken pox "Not the least sick"
		54 m 20 d	About over the chicken pox
55 m 12 d	Coughing badly for a couple of days. Ran out without hat or coat Tonsils larger than before, especially right.	55 m 13 d	Still has some pox on the face
56 m 9 d	Tonsils and slight adenoid tissue removed		
61 m. 24 d.	Vaccinated again.		
62 m 4 d	High fever for two days, headache and arm pain.	62 m 15 d	Mumps, both sides, high temperature, Only took liquid food
65 m	Coughs terribly, nose and lips red		
65 m 17 d	Pronounced whooping cough. Given two doses of antitoxin. Bad coughing spells but no real whooping		
66 m 19 d	Quarantine sign lifted Several coughing spells daily		
		76 m	Tonsils removed
		77 m	Positive Schick. Three immunizations
81 m	Measles More severe than *B*'s attack a week or two later Eyes affected, requiring lancing of abscess Had to get glasses	87 m.	Pink eye in right eye Much complaint of head after eye cleared
		88 m 12 d	Influenza with high temperature.
		88 m 18 d	Still coughs at night Abscessed tooth.
		94 m 18 d	Cold with coughing
		94 m 27 d	Infantile paralysis
105 m	Chicken pox Had more pox than *B*		

TABLE 9 (*continued*)

Age	Subject *A*	Age	Subject *B*
	Illnesses (continued)		
114 m	Attack of influenza lasting one and a half weeks		
	Falls and injuries		
		6 m 2 d	Fell head first from bed, but did not show a scratch
7 m 15 d	Fell off bed Dreadfully frightened	7 m 0 d	Fell head first from bed
8 m 22 d	Fell from lap head foremost on edge of stone step Cried considerably Perhaps screen door broke fall Apparently not much hurt	8 m 15 d	Fell from baby buggy to concrete floor No bumps or scratches, but screamed quite a while
8 m 23 d	Fell on floor from high chair, landing on side Apparently not seriously hurt		
9 m 1 d	Fell forward in high chair, frightened but apparently not hurt		
9 m 6 d	Fell from go cart		
9 m 21 d	Fell from baby buggy and bruised nose		
9 m 23 d	Turned baby walker upside down, hit on forehead, splitting skin		
11 m 6 d	Frequent falls Tumbled off porch, head striking soft dirt Slipped off porch with one foot and plunged down		
11 m. 20 d	Fell backward from chair The crosspiece hit her mouth and injured her gums which bled profusely Apparently did not loosen teeth, but may have caused recession of lower teeth		
13 m 9 d	Fell out of go cart, bruising left nostril.		

TABLE 9 (*continued*)

	Subject *A*		Subject *B*
	Falls and Injuries (*continued*)		
14 m 12 d	Fell from father's lap and cut forehead on corner of chair, retaining the scar for over a month		
		16 m	Fell and bit her tongue several days ago
		16 m 12 d	Slid off threshold on back porch
16 m 27 d	Fell down one flight of marble stairs from second floor. Much crying. No serious injury apparent		
19 m 10 d	Continues to fall and hurt herself		
		25 m 20 d	Fell off chair, cried for long time
		28 m 18 d	Fell off chair lighting on head about two weeks ago
34 m 10 d	Fell one story onto a cement floor when screen gave way, landing on the thigh	30 m 28 d	Hit accidentally under left eye by *P*'s head. The lower left lid puffed up. The eye became inflamed and discharging
87 m	Fell one story from ice covered porch through banister to frozen ground. Seriously shaken up with evidence of spinal wrenching		
144 m	Received serious head blow from a girl stealing up from behind	103 m 4 d	Girl pulled chair away, plunging *B* to the floor, the back of the head striking against the chair, dizziness and pain for a couple of days

TABLE 10

SCHOOL MARKS, BASED ON THE AVERAGE OF ALL SEMESTER OR YEAR-END AVERAGES FOR ALL BRANCHES SCORED AS FOLLOWS *A* (OR 90-100), 4 POINTS, *B* (80-89), 3 POINTS, *C* (70 TO 79), 2 POINTS, *D* (60-69), 1 POINT

Grade	Age	Subject *A* Average grade	Age	Subject B Average grade
1			6 y 6 m	4 00
2	7 y 9 m	3 70	7 y 6 m	3 75
3	8 y 9 m	3 92	8 y 6 m	3 5
4	9 y 9 m	3 70	9 y 6 m.	3 66
5	10 y 9 m	3 53	10 y 6 m	3 0
6	11 y 9 m	3 66	11 y 6 m	3 10
7B	12 y 4 m.	3 12	12 y 1 m	2 33
7A	12 y 9 m	3 25	12 y 6 m	2 81
8	13 y 9 m	3 57	13 y 6 m	3 88
9	14 y 9 m	3 25	14 y 6 m	3 50
10	15 y 9 m	4 00	15 y 6 m	3 60
11	16 y 9 m	3 80	16 y 6 m	3 66
12B			17 y 1 m	3 50
12A			17 y 6 m	3 75
College Freshman	17 y 5 m	2.16	18 y 1 m	2 66
College Freshman	17 y 9 m	2 16	18 y. 6 m	2 33

Y, year, m, month

above in the former case and of 16 and above in the latter case, the data have throughout been computed on the basis of both divisors. Although Terman recommends the use of 16, most Binetists probably use 14 The basis of computation for ages above thirteen has been greatly modified in the new Terman-Merrill scale—whether for good or evil is not now the question. (The need of a corrective formula, it is interesting to observe, was foreseen by the writer after a brief adventure with the Binet *IQ* in 1916) (4) Likewise when the results are based on the series of annual testings—i e, the tests that were given on or near the subjects' birthdays—the *IQ*'s are almost identically the same, with *B* retaining again a negligible superiority of 0 6 *IQ* when the divisor used is 14 and *A* retaining a negligible superiority of 0 3 *IQ* when 16 is used

Is this amazing similarity, not to say identity of results, covering a period of from 13 to 14 years of Stanford-Binet testing due to the sameness of the original germ plasm or of the external physical and social environmental stimulations or to an equivalent concurrent

action of both variables? While these subjects grew up in the same parental homes except for the absence of B during the first four years of A's life, and in the same cities, that does not necessarily insure the same psychic climate or the same subjective experiences. Moreover, all of the eight schools attended by A during the testing period were different from the seven schools attended by B, with one exception. Only three of the schools were parts of the same school systems. Of their numerous teachers only one came into direct teaching contact with both girls. In spite of the vicissitudes in school and teacher backgrounds, whatever the differences may have been, the average *IQ*'s for the entire series of tests are practically identical.

Does the similarity of the *IQ*'s in spite of the differences in the school backgrounds constitute an argument in favor of the dominant rôle of native endowment so far as intelligence is concerned, or of the dominance of the joint rôle of gene heredity and home environment? Does it indicate that the differences in the schools were negligible as compared with the joint force of heredity and home backgrounds? Or were the differences between the schools so slight as to be negligible?

Much the same situation obtains with respect to comparative variability as determined by two *SD* measures, the *MV*, the *Q*, and the *PE* computed from the *IQ*'s. In the annual series of tests the variability is slightly larger for B in all the measures save one, but the differences are small, varying in terms of the two *SD* formulas from 0.20 to 0.25. In the semi-annual series the variability is somewhat larger for A when 14 is used as the divisor and for B when 16 is so used. The difference, based on the two sigma measures, varies from 0.53 to 0.64. In terms of the coefficient of variability based on the *SD* the differences between the two testees are unimportant. They amount for the annual series to .001 when age 14 is the divisor and .002 when 16 is so used, while the corresponding figures are .004 and .006 for the semi-annual series.

How does the variability of these subjects compare with prior findings? The average *PE* of five earlier studies of retests of children of various ability levels tested at varying intervals is 5.8, which is perceptibly higher than the *PE*'s for our subjects (1, 6). In the semi-annual series, the *PE* for A is 3.07 when age 16 is used as the divisor and 4.31 when age 14 is the divisor. The corresponding

PE's for *B* are 3.50 and 3.95. The figures for the annual series of tests are about the same. Based on the most frequently used *SD* formula $\left(\sqrt{\frac{\Sigma(d)^2}{n}}\right)$, the variation is from 4.40 to 6.27 in the different series for *A* and from 4.65 to 6.05 for *B*. While the variability does not differ greatly for the different age periods, from 3 to 7, 8 to 12, and 13 to 16 or 17, yet it is slightly greater for both subjects for the period 8 to 12 (based on both *SD*'s and the *MV*'s) and perceptibly greater for the period 13 to 17 when age 14 is used as the divisor but not when age 16 is so used (Table 7).

In making comparisons with other investigations, it should not be forgotten that our results are based on at least 26 measurements while most other findings are derived from only two or three retests. Perhaps the normal expectation from frequent retesting would be a reduction of variability.

In spite of the relatively small amount of average variability for the entire series, the difference between the extreme *IQ*'s for these two subjects is very considerable. The range between the extreme *IQ*'s is about as great as many ranges previously reported for normal and deviating subjects. Conclusions based upon either extreme would not have given a correct picture based on the means for the series. The range for both of our siblings amounts to 27 *IQ*'s, from 106 at age 3 to 133 at 15½ in the case of *A*, and from 105 at age 11 to 132 at age 17 in the case of *B*. However, the highest *IQ*'s for both subjects may be regarded as spurious because they are based on five passes in age XVIII, giving a mental rating of 18-6 from the XVI-year base instead of 18 had all the 18-year tests been passed.[3] Disregarding these spurious *IQ*'s, due to scoring imperfections in the scale, *A*'s range still amounts to 23 *IQ* points (to 129 at age 16 based on 14 as the divisor) and *B*'s to 20 (to 125 at age 17-6 based on 14 as the divisor). If we confine the comparisons to the ages below 15 which are not complicated by the question as to which divisor to employ, the range still reaches 17 points for *A* and 18 points for *B*, which is sufficient to jump the classification from average to superior intelligence in the generally followed ability classification.

[3]As explained in Table 1. Nevertheless, *B* obtained a still higher *IQ* in Form *L* of the new Terman-Merrill Binet revision, namely, 136 at 17-8.

2. *Differences in Test Findings*

But the test performance of these siblings is not alike in all respects. A detailed analysis of the results presents strikingly curious differences in the two Binetgraphs. An interesting spectacle is presented of fluctuations or irregularities of mental development that occur at different age levels. In the earlier part of the century the literature on mental development was replete with references to plateaus and spurts in the curve of mental growth with little reference, however, to actual losses or deteriorations in the case of normal children. More recently some writers have concluded from Binet measurements that these assumptions are mythical. They affirm that the curve of mental growth is essentially smooth or uniform in character. But our Binet data show that *A* suffered losses in *IQ* in 10 ages in the semi-annual series, varying from 2 to 9 *IQ* points. The losses average 5.1 when 14 is used as the divisor and 5.8 when 16 is so used. In the case of *B*, 13 losses occur (in two cases the comparisons are based on the preceding year because of gaps in the series), ranging from 1 to 8 *IQ* points but the average loss is only 3.7 based on 14 as the divisor. While the losses do not differ greatly for the two subjects, the largest and smallest losses occur at different ages. The largest declines occur for *A* in ages 8½ and 11½ (9 *IQ*'s) and for *B* in age 14 as compared with age 13 (8 *IQ* points). The smallest losses occur for *A* at 5½ and 13 (2 *IQ*'s) and for *B* at 7½, 12, and 15½ (1 *IQ*). The losses for the two subjects occur at corresponding ages in only 7 ages, 4 to 7½ and ages 15 and 16. Apparently the losses cannot be due solely to certain peculiarities in the measuring scale.

The drops in the Binet *IQ* are not paralleled in the case of either subject by appreciable losses in the same ages in the Seguin curve except in one or two ages.

On the basis of the annual birthday testings, *A* suffered 6 losses varying from 1 to 7 *IQ*'s and averaging 3.3 *IQ*'s, while *B* sustained 7 losses, varying from 2 to 8 *IQ* points and averaging 5.4 *IQ*'s. The losses occur at corresponding ages for the two subjects in only 3 instances, ages 6, 14, and 16. *A* suffered a consistent decline from an *IQ* of 119 at 6 to 110 at 9, a drop of 9 *IQ* points and *B* declined from an *IQ* of 119 at 9 to 110 at 14, with one exception at age 13 (where the *IQ* is 118), the same amount of decline as for *A*.

With respect to losses in intelligence age level, curiously four such instances occur in the semi-annual series in *A*'s record, varying from two months to six months (average 4.2 months) while only one loss of one month occurs in *B*'s record, although there are two instances of no gain. *B* suffered more losses in *IQ*'s in the semi-annual series, but the losses averaged less than for *A*. None of the losses or failures to gain occurs in corresponding ages for the two subjects, again suggesting that some factors extraneous to the tests themselves were operative. In the annual series of testings at the birthdays no losses occur although the Binet levels for ages 13 and 14 are the same for *B*.

Pronounced gains occur in the series of semi-annual tests for *A* between ages 11½ and 12, amounting to 10 *IQ*'s, between 3 and 3½, 8 *IQ*'s, and between 4 and 4½, 7 *IQ*'s, and for *B* between 6½ and 7, 12 *IQ*'s, between 14 and 14½, 9 *IQ*'s and between 3 and 3½, and 11 and 11½, 7 *IQ*'s. Thus the spurts for the two subjects also do not occur at corresponding ages, except between 3 and 3½. In the annual series of birthday tests, the pronounced increases occur for *A* between 4 and 5, 12 *IQ*'s, and between 15 and 16 when 14 is used as the divisor, 7 *IQ*'s, and for *B* between 14 and 15, using 14 as the divisor, 10 *IQ*'s[1] and between 6 and 7 and 12 and 13, 7 *IQ*'s. Here again the spurts occur at different chronological age levels for the two subjects. In this series *A* shows a fairly consistent gain in *IQ* from age 9 to 16, while *B* shows no such consistent gain except from 14 to 17 when 14 is used as the divisor. The Seguin record is not sufficiently complete to permit comparisons for all these ages but a parallelism is apparent in at least two or three ages.

Based on the annual series, the largest gains in Binet ages occur in *A*'s record between 4 and 5, amounting to 20 months, between ages 10 and 11, 19 months, and between 12 and 13, 18 months; and in *B*'s record between 12 and 13, 24 months, between 11 and 12, 21 months, and between 6 and 7, 20 months. The stages of greatest increase in terms of *IQ*'s do not exactly correspond to the periods of greatest increase in terms of months, but the differences are not important.

Before seeking an explanation for these discordancies in the two intelligence growth curves as determined by Binet measurements,

---

[1]Ages 16 to 17 are not included because of the spurious increments at 17

let us direct attention to an even more curious phenomenon of see-sawing in the two records. In the series of semi-annual tests, *A*'s *IQ*'s are higher than *B*'s in 13 ages, lower in 11 ages, and equal in one age—age 8. *A*'s superiority varies from 2 to 12 points and *B*'s from 1 to 14. But the most anomalous phenomenon in the shifts is that *B* is consistently superior to *A* in Binet performance in the early stages, then *A* forges ahead, then *B* takes the lead, and finally *A* comes out in front and retains her position of leadership. To cite the essential facts: Before the age of three, in which some of the tests from the 1905 and the Stanford-Binet were employed without any attempt to compute *IQ*'s, *B*'s superiority is barely apparent. But from ages 3 to 4½ in the semi-annual series she excels at every testing period, the average superiority amounting to 8.8 *IQ*'s, which is a very marked difference for the post-infant stage of development (Table 5). On the other hand, from ages five to 6½ *A* is ahead by an average of 4.0 *IQ*'s. From 7 to 10 *B* is again in the lead except at age eight. Her average superiority for the five ages in which she excels is 6.5 *IQ*'s. But from 10½ to 16, *A* is ahead in all ages which can be compared except at 11½. Her average superiority for the ages in which she excels from 10½ to 14 (exclusive of the ages that are complicated by the question as to the divisor to use in computing the *IQ*) amounts to 5.1 *IQ*'s.

In harmony with these findings, it is interesting to note that the average *IQ* of *A* tends to increase with each successive period while that of *B* tends to diminish, with one exception for each subject. Thus the increase of *A*'s average amounts to 6.5 *IQ*'s from the period 3 to 4½ to the period 5 to 6½ (Table 5), it decreases 5.1 from the latter period to the period 7 to 10, and then increases 8 *IQ*'s from the latter period to the period 10½ to 16 based on age 14 as the divisor and 3.8 based on age 16 as the divisor. On the other hand, *B* drops 6.3 *IQ*'s from the first to the second period, then increases 5.4 *IQ*'s from the second to the third period, and then drops 4 *IQ*'s from the third to the fourth period based on age 14, and 7.8 based on age 16. Apparently if practice from repeated retesting, at six months' intervals, affects performance in the Binet tests, the effect of such practice is not uniform for these two subjects, unless the assumption is entertained that other factors may have distorted the practice effects. It has been affirmed that Stanford-Binet "tests taken a year apart show no effect of practice" (1). But

we cannot categorically conclude from this fact that practice is negative when the interval between the tests is shortened 50 per cent.

With respect to these anomalous shifts in the relative ability of the two subjects and the stagnations, retardations, and accelerations manifest in both Binet records, two questions merit careful consideration. First, what about the validity of the apparent shifts in the relative ability of the two siblings during the successive life periods? Is there any corroborative evidence that they are genuine? Is the shifting status paralleled in the psychomotor tests or in the developmental schedules based on the observation of behavior reactions? Second, do the developmental data or other available criteria explain the asymmetries in the two Binet curves of development?

So far as concerns the results of the psychomotor tests, the correlation between the fluctuations of the Binet performance of the two subjects and their peg board, Seguin, and Healy *A* and *B* performances seems to be rather negative or at least ambiguous. Thus from the 12th to the 54th month (4½ years) in which *B* excels in the Binet (slightly before age three and decidedly after age three) *A* excels in the peg boards (Table 3). She does better at 21 months than *B* does at 24 months in four comparisons and is inferior in only two. At 30 months, *B* has surpassed *A*'s record at 21 months, although the superiority is not very patent in the more difficult boards. At 54 months *A*'s record surpasses that of *B*'s at the same age in four of the comparisons while *B* excels in only three, although *B* had done the test seven times at that age as compared with only three performances for *A*. From age two to 4½ *A* is consistently ahead in the Seguin, her performance at two being superior to *B*'s at 2½ and her record at three being superior to *B*'s at 3½. Her average superiority in the Seguin for the years from three to 4½ amounts to 11.3 seconds based on the best of three trials and 24.6 seconds based on the average of three trials (Table 6). In the possible comparisons at the same ages in the interval from 5 to 6½ in which *A* is ahead in the Binet, *B* is "out in front" in the Seguin. In the interval from 7 to 10, however, in which *B* rates higher on the Binet, *A* scores higher in the Seguin in six of the comparisons and lower in only three. But the average differences for the entire age interval are trivial. Based on the averages of all the tests given during this period *A* excels by only 0.3 seconds based on the fastest responses and by 0.91 seconds based on the averages. During the

above periods the Seguin test was administered 13 times to each subject. During this period, however, we meet with another apparent discrepancy in that *B* excels rather patently in Healy *A* and Healy *B*. Although accidental successes occur for *A* at 4½ and for *B* at 5, *B* required 29 moves and 114 seconds to do Healy *A* at age 8 the first time and 15 moves and 55 seconds the second time, while *A* at the age of 8½ required 69 moves in two trials during a total of 256 seconds. She required 201 seconds and 52 moves to do Healy *B*, while *B* at the same age required 87 seconds and 23 moves. These two tests showed some conformity with expectation. For the following period, from 10½ to 16, the performance tests were not administered sufficiently frequently at corresponding ages to permit comparison.

With respect to the record of growth and behavior characteristics summarized in Table 8, which may be regarded as indicators of growth in intellectual maturity or developing ability, it is necessary to point out that the data included in the tabulation are highly condensed from the mass of available observations which necessarily cannot be reproduced inextenso. It is believed, however, that the most salient items are presented in the table. It should also be remarked that the observational data are largely confined to the first few years of life.

These data, taken all in all, seem clearly to corroborate the Binet verdict regarding *B*'s superiority during early childhood (although, as already stated, the superiority during the infantile stage is not so marked), as against the verdict of the psychomotor tests, although flaws exist in the observed facts. Some of these flaws may be due to gaps in the records, as neither child was under constant observation, but some of them are doubtlessly due to other factors.

Reference will first be made to certain behavior characteristics in which the data are somewhat ambiguous or discrepant.

Thus, although *A* made propulsive, walking movements when held up several months before *B* did, *B* was a few days ahead in actually taking a few steps unsupported. Perhaps this anomaly can be explained by the circumstances that *A* was hampered at the outset in learning to walk. She started to walk on her toes and thus readily lost her balance. When she overcame this initial handicap she made very rapid progress, so that she was able to run at about the same age as *B*. She soon became very fleet-footed and a very

active runabout who remained such for several years. She also became very clever in eluding her guardian Thus at 22 months of age she ran away several times in a small village and at 34 months she escaped from her guardian in Los Angeles and ran a half dozen blocks before being apprehended in the middle of the street by a passerby who took her in tow and notified the police department In Denver six weeks later she escaped from the house and had penetrated a considerable distance into a park near-by before she could be tracked down.

Similarly, although *A* was slightly ahead in trying to raise herself to a sitting position, *B* succeeded in doing so about a month earlier *B* was almost a month ahead in the ability to sit on the floor unsupported for a brief interval *B* was ahead in crawling, for *A* never crawled Nevertheless she began to roll in bed about two months before *B* began to crawl on the floor and she became as agile in rolling on the floor to any desired objective as she later became adroit in running about. On the other hand, while *A* was able to stand on the floor by holding to a support about three weeks earlier than *B*, *B* stood momentarily unsupported about seven weeks earlier than *A*. Contrariwise, while the smiling reflex or a response reaction resembling a smile, whatever its significance, appeared in *B* about three weeks earlier than in *A*, *A* laughed earlier (by three days) and chuckled earlier (by two days) than *B* did. So far as fingering, grasping, and handling objects are concerned, *A* surpassed *B*. At three months she fingered a book and bath tub faucet, a week later she made rudimentary attempts at hitting a rattle and at grasping it, at a little less than four months she was grasping objects near-by and carrying them to the mouth, and a week later she was holding the nursery bottle with both hands. *B*, on the other hand, did not begin to show signs of grasping until she was four months and nine days old, or to hold the nursing bottle until she was four months and 24 days old. *A*'s early ability to grasp and handle finds a parallel in her later skill in peg board and Seguin performance

We turn now to the behavior responses in which *B* seems to possess evident superiority in varying degree The list includes the following items·

Getting up from the floor unassisted, by 15 days.

Getting up on a settee, a rocker, or Morris chair from a small stool, by about a month and a half

Pulling herself up on a swivel chair or big rocker, by slightly more than a month

Indulging in prespeech chatter or babbling, by about a month

Reproduction of speech sounds (*da da*), by a month

Meaningful expressions or short phrases, by over five weeks

Speech imitations of sounds of animals or things, by a month and a half

Early manifestation of interest or concentration of attention, by 11 days

Noticing her hands, by three weeks

The tendency of the eyes to follow the form of the moving parent, by six days

Turning the eyes and head toward the moving parent, by one month

Beginning to give more attention to sounds, about 2½ months

Opening the door for the first time, about 8 months

In two things *A* is clearly in the lead during this infantile and post infantile period, namely in the age of appearance of the deciduous teeth and in reading ability. *A*'s advantage amounts to 2 months and 26 days in the case of the lower central incisors, 5 months for the left upper incisor, 3 months and 11 days for the upper right incisor, and one month and 27 days for the upper canines. As compared with the norms, *A*'s eruptions occurred rather early and *B*'s rather late during the period in which *B*'s development seems to be accelerated in many other respects. Here is another anomaly in the two records. Why should *A* forge ahead in dental development during the period in which *B*'s development was ahead in most other respects, particularly in view of the fact that *B* received more attention from pediatricians and did not experience the dietary difficulties that *A* did? It has been held that a correspondence exists between dentition and carpal development and that carpal development affords a reliable index of both physiological and mental maturity. Our siblings thus present evidences of discrepancy between the stages of anatomical and psychological maturity.[5]

In the matter of reading proficiency *A* was markedly superior to *B* in early life. At 16 months and 28 days *A* knew the letters

[5]It is interesting to observe, however, that during the fourth period, 10½ to 16, when *A* forged ahead in the Binet, menstruation began for *A* at eleven and for *B* at about 14. In this particular, correspondence seems to exist between physiological and mental maturity.

B and *R* from playing with alphabet blocks At 17 months and 19 days she recognized *B, R, O, I,* and *P* At 37 months and 25 days she had no difficulty in matching word cards, and soon learned the words *yellow, little, green, can, red, fly, has, in, we, dog, kitty,* and many others She also read two sentences correctly "*The little girl can run,*" and "*The boy can run*" She progressed rapidly in reading without receiving much instruction and at 42 months read two short stories about "*Tom's Dog*" and a "*Fat Cat.*" The record at 42 months, however, indicates that she had shown little interest in reading for some time and that the instruction had been discontinued. During the corresponding stage *B* was wholly incompetent in reading, but her incompetency was no doubt due purely to lack of environmental stimulation for no effort was made to teach her to recognize letters or words. Obviously *A*'s reading superiority is attributable wholly to the factor of experience or learning

Unfortunately most of the above comparative data from tests and observations are limited to the first few years of life, particularly the first two when the difference in the Binet is less marked than during later years Few systematically recorded data for the later years are available except school marks, and these are missing for age six for *A* As everyone knows, school grades are not very reliable criteria of either intellectual ability or scholastic proficiency, especially when they are issued by many different teachers whose standards of marking may differ markedly With this reservation in mind we have supplied in tabular form (Table 10) the average marks of all the teachers for each successive semester or grade in terms of arithmetical points Four points were arbitrarily assigned to a grade of *A* (or its equivalent), three points to a grade of *B,* two points to a grade of *C,* and one point to a grade of *D*

The difference in school marks is not striking. The evidence from this source is rather equivocal *A* possesses a slight advantage throughout her school career except at the college level. Thus she has a slight edge on *B* from 7 to 10, the period in which *B* excels in Binet age From 10½, to 16, the period in which *A* has higher *IQ*'s, she has higher marks in six grades and lower marks in only two grades. On the other hand, her marks are lower than *B*'s for both college Freshman semesters As a college Freshman, however, she was over a half year younger than *B,* while in the public schools she was about three months older than *B* in each grade The differ-

ence in maturity somewhat complicates the results Moreover, *A* graduated from an eleven-year and *B* from a twelve-year public school system

D Explanation of Fluctuations in Binet *IQ*'s and Binet Levels

In attempting to discover the causes for the fluctuations in the Binet curves—the factors responsible for the retardations and accelerations—one must consider both endogenous and exogenous factors So far as retardations or losses in *IQ* are concerned one might look for resistances, emotional upsets, restlessness, inattentiveness, distraction, fear of the examiner, dislike for the tests, lack of intellectual stimulation in the subject's ordinary social and intellectual environment, illnesses, and injuries that might affect interest, attitude, and the output of psychophysical energy, and defects in the tests used so far as concerns their construction, standardization, and method of scoring and administration Indeed it is conceivable that arrests or even deteriorations in development might be primarily occasioned by ill-defined hereditary or constitutional factors. Irregularities in the strength of the hereditary propulsion, the growth impulse, or constitutional drives have perhaps more frequently been suggested in explanation of spurts of development Other possible explanations of accelerations in intellectual growth would include improved physical health, a stimulating social and educational environment, freedom from hampering conflicts, fears, and anxieties, from antipathy toward the examiner or the testing procedure, and unsatisfactory construction, standardization, or methods of scoring the test used It is obvious, of course, that defects in the measuring instrument might produce spurious fluctuations in both directions

Are there any data at hand that enable us adequately to explain at least the more marked fluctuations that occur in the records of these two examinees? Let us turn first to possible explanations of the stagnations or losses in Binet level or Binet *IQ*

1 *Causes of Losses*

In Table 9 is chronicled in parallel columns, the records of colds, influenza, childhood infections, other illnesses, falls, and injuries A minute analysis of the data permits only very guarded or tentative conclusions To illustrate A month after *A*'s whooping cough,

which occurred at 5-5 (five years, five months), the Binet test shows a loss of two *IQ*'s. *B*'s milder attack at 14 months shows no clearly discernible effects. Four months after an attack of measles at 6-9 *A* shows a gain of 2 *IQ*'s, while *B* after a milder attack at 30 months apparently suffered no injury. One month after an attack of chicken pox at 8-9, *A* shows a gain of 2 *IQ*'s from a low score of 108 *IQ*'s followed by another gain of 2 *IQ*'s six months later, although she had had an attack of influenza lasting a week and a half two weeks before the latter test and was very restless during the test. On the other hand, *B*'s *IQ* obtained at the time she had the chicken pox at 4-6 fell five points. Eight months later she lost four more points after an attack of mumps. One year and two months later at the age of 6-4 she had her tonsils removed. The Binet test two months later yielded the lowest *IQ*, 109, up to that time. However, eight months after the operation the *IQ* had advanced to 121, the highest point reached up to that time except at 3-6 (124 *IQ*) and her *IQ* remained fairly high until the age of 9. On the other hand, while the removal of *A*'s tonsils at 4-8 was followed by a gain of 5 *IQ*'s in the test four months later, the gain was followed by a slow recession that continued for several years with several fluctuations and that reached a low point of 108 at 8-6 when she complained of being ill during the examination so that the test had to be given in two sittings. The subsequent *IQ*'s did not reach her five-year level until she reached 12-6. At the age of 7-10, *B* suffered a mild attack of infantile paralysis leaving no motor sequelae but a certain amount of continuing emotional instability. While her *IQ* a month after the attack had dropped three points, six months later it had reached the highest point, 122, since the age of 3-6 when it was 124. These results do not supply incontestable evidence either for or against the hypothesis that Binet retardations can be explained by fluctuations in the physical condition of the subjects. The possibility that the fluctuations might be due to the coarseness of the measuring instrument should be given due consideration before too much emphasis is placed on factors extraneous to the scale.

What does the record of physical injuries show? Most of the falls and injuries suffered by both subjects occurred in early life and were minor ones so far as can be inferred from the record. Of course, it is recognized that accidents that appear to be trivial from the standpoint of the observable physical injuries, may be very

serious from the standpoint of the psychic trauma produced, of which the observer may be unaware, which may continue subtly to color the individual's mental reactions for years. *A* suffered a severe jolt at 34 months from hurtling through a window when the screen became loosened. She plunged one story to the concrete pavement below, landing on her thigh. Her Binet *IQ* two months later was the lowest in her entire series of tests. One cannot conclude unqualifiedly, however, that this physical shock was solely responsible for the low *IQ*. Her low *IQ* in early life may have been the resultant of the cumulative effect of numerous falls. During the first year she had a record of 10 falls as compared with three for *B*. During the first 16 months the corresponding figures are 14 and 6, and during the first three years 15 and nine. On the other hand, the record of colds, fevers, childhood infections, and illnesses does not differ greatly for the two subjects. During the first year each sibling has a record of six infections or illnesses, during the second year *A* has four and *B* has seven, during the third year they have two each, and during each of the two following years one each. At the age of 7-3, *A* had another severe fall, sliding from an ice covered second story porch to the frozen ground below, and suffering a severe wrench of the spinal column and nervous shock. The *IQ* three months later showed a drop of 4 *IQ* points, followed by a rise of 6 *IQ*'s six months later. This was followed by a greater drop a year later but this might have been due to the fact that she was feeling ill and had to be tested in two sittings, as already explained. In spite of *A*'s numerous falls, however, her average *IQ* for the entire series is about the same as *B*'s, while her average *IQ* for the final period is higher. *B* suffered one head injury worth mentioning. At the age of 7-4 she was precipitated to the floor when a girl from behind pulled her chair away as she was about to sit down. In the fall her head struck against the edge of the chair. For some time she complained of dizziness and pain in the head. Her *IQ* five months later fell 3 *IQ*'s and continued to fall for five years, with possibly one exception. Perhaps the cerebral injury or the psychic trauma produced by the accident may have had something to do with the *IQ* recession, but the evidence on this point is not incontestable. For the greatest losses, which occur for *A* at 8½, 9, and 11½, and for *B* at 6½, 11, and 14, based on both the semi-annual and the annual series, the records fail to disclose any

explanatory data except the illness of A at 8-6 and her chicken pox at 9

The supposition is probably correct that inattention and restlessness during the test constitute one of the endogenous sources of IQ losses. But our records show that this result does not necessarily follow The evidence is discrepant on this point A was restless and inattentive during the early tests while B was attentive and responsive This may be a factor in B's slight superiority during this period At the age of 3-0 A was restless during the test and scored her lowest IQ At 3-6 she exercised better control and made an IQ of 114 But at 4-0, although she gave good attention and was not restless, her IQ fell to 108 On the other hand, at 4-6, although she was constantly fumbling with her hands and her attention was constantly wandering, her IQ rose to 115 At 6-6 she was fidgety all through the test, indulging in excessive finger activity, and her IQ dropped to 113 At 7-6 she was again restless and the IQ fell to 111, but three months earlier she had suffered one of her serious falls Likewise B at 3-6 was very restless and impatient to get to her kitty, but in that test she secured one of her highest IQ's, 124 In later adolescence she became rather resistant to the testing ordeal, partly because she had grown rather bored with the tests and partly because the testing interfered with her interest in other pursuits This apparent apathy may have been a factor in her declining IQ's during part of this period but it would be speculative to assert that the antithetical condition fully accounts for the spurts that occurred

2 *Causes of Spurts*

The attempt to explain the numerous spurts that occur in the two curves is even more difficult than the attempt to account for the losses To cite a few specific situations There is nothing in the record to explain A's large IQ spurts from 11-6 to 12 (amounting to 10 IQ's), from 3 to 3-6 (amounting to 8 IQ's), or from 4½ to 6, unless it is the removal of the tonsils at 4-8 But the spurt apparently started at 4-6 before the tonsils were removed The literature on the results of the orthogenic effects of the removal of physical handicaps teems with contradictory findings Even granting that B's large gain from 6-6 to 7 (12 IQ's) is explained by the tonsillectomy, there is nothing in the record to account for

the large gains from 11 to 11½ (7 *IQ*'s) or from 12 to 13 (7 *IQ*'s or 24 months). Since the spurts occur at different ages for the two subjects, just as do the losses, it would appear to be rather gratuitous to attribute them to hypothetical changes in the stimulation of the home environment. While the home environment may have been physically the same for the two subjects during the period of the tests, it is, of course, admitted that the psychic appeal may have been vastly different for the two subjects. There is, however, no means of accurately estimating each respondent's subjective reactions to the home environment. There is no means of measuring the difference in the stimulating or repressive effect that the home life exerted on the intellectual life of each sibling. The response that anyone makes to environmental stimuli depends not only on the nature of the stimuli, but also on the individual's temperament, attitudes, habits, mind-sets, drives, purposes, and ideals.

That there were marked differences in the school environments of the siblings throughout the school period is probable. That these differences may have some bearing on the test fluctuations is possible. But it would lead us into a purely speculative adventure to attempt to determine to what extent the rises and falls in the two curves of intelligence are due to differences in the efficiency or the stimulation value of the different teachers and classrooms that affected the lives of these youngsters. As we have seen, it is the contention of one of the recent researches that the *IQ* can be radically improved, and that the chief factor in this transformation is a superior school environment. Be this as it may, we have at hand no reliable yardstick for measuring the relative efficiency of the schools that influenced the intellectual growth of these subjects. Although they probably rated better than average, we have no reliable criterion by which to appraise the contribution each school made to the development of the mental vigor of these two pupils. A priori we might conclude that, all in all, they were equally efficient because the final product was two young ladies with *IQ*'s that did not diverge markedly. But would the levelling processes have been just the same if the siblings had started out with markedly different *IQ*'s? According to our traditional concepts, patently no. In spite of the hypothetical similarity of environmental opportunity, however important or unimportant this factor may have been, marked fluctuations in the *IQ* curve did occur at non-corresponding points. But,

let us turn from these hypotheses, which may be true but cannot be substantiated from documentary data, to the consideration of another possible explanation of the fluctuations in the Binet curve that has been relatively disregarded in the literature, namely imperfections in the scale itself

E. Imperfections in Binet Test Construction and Scoring as a Source of *IQ* Fluctuations

Let us remind ourselves that the Stanford-Binet scale contains not only so-called power tests, but also tests of information or erudition the successful performance of which is dependent solely or largely upon experience or learning It is obvious therefore that the results of instruction will be reflected in Binet performance Again, nearly all of the tests are scored on the all-or-none principle although the responses to some tests cannot fairly be scored as all wrong or all right. Many test situations admit of borderline answers, answers that are neither wholly correct nor wholly wrong, but correct in varying degree In consequence, it is sometimes very difficult to determine whether a given response should be scored plus or minus, because it may be correct or incorrect in accordance with one's point of view, or it may be rather more correct than incorrect, or slightly more incorrect than correct It's a toss-up whether to score such responses plus or minus Different examiners will inevitably score uncertain responses differently How important the factor of scoring discrepancies is may be shown by the fact that learners in Stanford-Binet testing have differed by 30 *IQ* points in scoring the same test sheets But no matter how well trained or experienced the Binet examiner may be, he is always confronted with the problem of scoring borderline responses. Now, it is apparent that if the examinee is marked minus on borderline responses that are very near the passing standard a later test may produce a spurious spurt, for he would be credited with an advance of from two to six months (or even nine months in the abbreviated scale) depending on the location of such tests, although the progress made may be very slight and would not represent a gain of from two to six months. Owing to the coarseness of the unit of measurement, the passing of only one test in the lower part of the scale means an increase of 3 or 4 *IQ* points The passing of one test at age fourteen scoring six months yields 4 *IQ* points. At this level a child who just falls

short of passing two tests at the first examination and who just passes them on the second examination receives an increase of 8 *IQ* points Such an increase, although it "makes the grade," yields an increment that is largely spurious Spurious losses in *IQ*'s would, of course, be explained by the converse situation, namely, marking a test as minus that practically coincides with the passing standard. Records with numerous borderline responses of this nature might thus contain rather extensive spurious gains or losses that would make a fairly uniform growth curve appear deceptively irregular

Another possible explanation of spurious fluctuations in the Binet *IQ* curve relates to the extent of the testing. If the testing is not equally complete during each successive test the losses and gains might be due to the inequalities in the range of the testing The importance of wide range Binet testing is shown by an experiment in 1916 in which the 1908 Binet scale was used with children of various intelligence levels When a wide-range method of Binet testing was employed, in which the number of advance tests given averaged 4.5 tests more than the number given in a narrow range method, used on an unselected group, "the average amount of extra credit earned amounted to three-fourths of a year or more" in most of the Binet ages. "In one age it amounted to a year and a half" (5)

In the same connection, the inadvertent omission of tests might readily become a factor irrespective of whether the narrow-range or the wide-range method of test administration is employed. Many instances of such inadvertence have been discovered in the inspection of the record blanks of many examiners in our clinic files Fortunately only one or two instances of such inadvertent omissions were found on the reexamination of the records of our two subjects. Recessions, stagnations, or accelerations in the curve due purely to differences in the thoroughness of the testing are, of course, fallacious

It would be impossible in this paper for obvious reasons minutely to analyze all the responses in all contiguous Binet age levels. All we can venture to undertake is to analyze very briefly the salient test data bearing on the theory of spurious fluctuations from imperfections of test construction or scoring based on the tests in and adjacent to the ages showing the largest losses and the largest gains in the Binet records

The large gain in *J*'s record of 20 months or 12 *IQ*'s between ages four and five is not adequately explained by errors of scoring due to borderline responses The plus score on counting four pennies in IV_3 may be too high because, although the pennies were counted correctly three times, the subject persisted in saying there were three pennies But this may have been done in a spirit of banter because she counted them correctly six months earlier The three commissions in V_6 were also scored plus which were performed correctly except that several boxes in addition to the keys were placed on the chair But this excess credit, if any there be, may be counteracted by the score given at four on test III_1 She first replied that she was a little boy Later, however, she said she was a little girl This confusion was probably due to the fact that she was wearing bloomers at the time and had been referring to herself as a boy She had passed the test six months earlier and no doubt knew whether she was a boy or a girl in spite of the ambiguous response The putative excess rating at age five may also be minimized or canceled by the fact that 50 per cent of the standard requirement in the comprehension test, VI_1 (What's the thing to do), had already been met at the age of four Her rating at 4½ might be slightly too high through another circumstance At this age she passed all the V-year tests and was thus given a V-year Binet rating according to standard practice, although she failed on IV_1, copying a square On the basis of a III-year base her *IQ* at 4½ would drop from 115 to 111 But this does not explain the high *IQ* of 120 at age five which is not affected by any failure below the base. The rating at this level may be slightly exaggerated because the examinee had passed 50 per cent of the comprehension test, VI_1 during the two preceding tests (she passed a different question on each occasion) The plus accorded the borderline answers to the definitions in $VIII_5$ would also tend to exaggerate the rating somewhat at age five However, they were also scored plus at 4½ based on somewhat dubious stereotyped descriptions.

The spurt at 4½ cannot be explained by the extent of the testing, for there were only 14 advance tests (i e , tests above the base) at that age as compared with 22 at age four. Nor can the spurt at five be thus explained for while there were 21 advance tests given at five as against only 14 at 4½, only one of the additional tests scored plus, the borderline vocabulary response.

The conclusion on the basis of the analysis of the test responses, is that the total amount of overscoring does not account for more than a minor part of the spurt at either age four or five and particularly the large increase at five

A suffered a loss of two months and 9 *IQ*'s at 8½ compared with age 8 although she was given 25 advance tests at 8½ and only 22 at 8 Possibly the rating at age eight is a trifle too high on $VIII_2$, counting backward, in which the time limit was exceeded by four seconds, and in IX_3, making change, which was scored plus although the answer to 10 minus 4 was first given as five and later as six On the other hand, two of the failures at 8½ perhaps do not represent any genuine loss of power The date test, IX_1, was failed so far as the day of the month is concerned (although it was passed at 8) and the counting backward, $VIII_2$, was also failed, because of errors, although it was scored plus at age eight (It was reproduced as 20, 19, 18, 17, 14, 12, 9, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 2, 1) These failures while real were probably due to slips of attention produced by the illness and head pains of which the testee complained during this test From these considerations the conclusions might be drawn that a part of the slump at 8½ was spurious while a part of it was transient, and so far forth fictitious The conclusion would then be that the deterioration at 8½ is largely spurious.

In the case of *B*, the greatest slump occurred at age 14 as compared with 13, a loss of 8 *IQ* with no gain in Binet level The failure on the XVI-year reversed digits test at age 14, which was passed at age 13, was counteracted by the passing at age 13 of XIV_8, three differences between a president and a king, on the basis of one correct distinction and one borderline distinction (namely, "the president gets in by being voted for and the king by custom " This was scored plus because at 12 it was explained that "the president is elected but the king's son follows the king ") Were this response scored minus, the 13-year *IQ* would drop from 118 to 115 but there still remains a five-point loss while the gain in Binet level amounts to only four months during the 12 months' interval If the vocabulary increase from 9,360 words at 13 to 10,620 words at 14 were reflected in a partial score, which it is not, although the gain obviously represents some measure of improvement, the loss would be reduced still further to a slight extent One might consider, however, that this theoretical loss is counterbalanced by a too

liberal credit at age 14 on XVI_3, abstract differences, which was scored plus on two correct responses on subtests and one borderline response and one of the subtests of which was correctly answered at 13. The advance of five months for this test appears to be somewhat excessive. All in all the analysis of the responses seems to show that at least part of the recession at 14 is genuine so far as the test finding is concerned The whole loss cannot be attributed to imperfections in the scoring techniques.

B's largest spurt is at age 7 compared with 6½, a gain of 12 *IQ* points and one year and five months Twenty-three advanced tests were administered at seven as against 18 at 6½ But the increment from the more extensive testing amounts to only two months, gained from a borderline response to the absurdities test in age *X* (The borderline response involves the third subtest: "he couldn't kill himself; somebody had to") Considering both facts, this gain on a generous construction might be construed as spurious Two other gains might be considered to be somewhat exaggerated by the method of scoring in view of the fact that 50 per cent of the test was passed at 6½ This includes VII_5, differences between concrete objects, and $VIII_5$, definitions The latter test in addition was passed on a borderline definition of balloon ("a round ball, red, all colors") One other possibility of excess credit at age seven derives from a favorable rating of a rather crude statement of similarity in $VIII_1$, namely between an apple and a peach ("they're red, both have stems") Counteracting in a measure these rather liberal ratings is a rather generous half credit score at age 6½ on the bow knot test which was scored minus at seven although the performance was very nearly as good as at 6½ After making all these allowances the conclusion would probably be that about 50 per cent of the gain at seven as determined by the Binet scale is genuine. Seven tests exclusive of the ten-year test were passed at age seven that were missed at 6½, while no test passed at 6½ was missed at 7 except the questionable bow knot performance. The basal age was six at both ages

To what extent borderline responses may invalidate or at least complicate the scoring and produce dubious fluctuations may be illustrated from a very brief analysis of the large upsurge between 12 and 13, amounting to two Binet ages and 7 *IQ* points. In the vocabulary test the passing of two additional words at age 13 (as

compared with age 12, yielding a score of 9,360 instead of 8,640) resulted in an increment of four months An increase of four months for only two words at the age of 13 seems like a spurious increment A similar increase at XII years would have yielded only three months and below 12 years only two months At age 14 the increase was 3½ words (to a score of 10,620) as compared with two words at 13, but no credit was received for this greater increase as the score fell short of the XVI-year standard (11,700). Incidentally, one more definition at the age of 16-6 would have yielded six months of additional credit and an XVIII-year rating in the vocabulary test It is obvious that such crudities of scale construction will produce deceptive irregularities in the curve Finally, attention may be called to the gain of five months at the age of 13 in the code test which allows two errors out of 11 words The amount of gain seems exaggerated in view of the fact that it was passed with one error as compared with four errors at 12, which fell merely two errors below the passing requirement.

The above analytic survey of individual test results, however brief, will suffice, it is believed, to render the hypothesis plausible that the extent of fluctuations in the Binet test scores from test to test are often exaggerated because of the equivocal or ambiguous nature of the responses or because of imperfections and uncertainties of scoring If this is correct, some of the conclusions reached regarding the extent of the transformation of the *IQ*'s in either direction produced by environmental stimuli are exaggerations Exaggerated increases in test ratings might be produced by the examiner's favorable anticipations Because of his favorable mind-set borderline responses are given the benefit of the doubt and accorded full credit On the other hand, an unfavorable attitude induced by unfavorable environmental situations might be unconsciously reflected in a hypercritical treatment of borderline responses, in consequence of which such responses are scored minus An unconscious attitudinal bias in either direction toward borderline responses might produce considerable shifts in the *IQ*'s in records containing many such responses. In reply to this assumption it might be argued that the effects of a good or poor environment are reflected not only in linguistic or verbal tests, such as the Binet, but also just as notably in performance tests, such as the Merrill-Palmer (as reported in the Iowa experiment) which are free from questionable responses The reply

is two-fold First, the scoring of performance tests may also present difficulties The scoring often cannot be done purely mechanically Many motor performances come within the borderline region of passing and the scoring can be tipped in the direction of the conscious or unconscious bias of the examiner. Second, the results of experimental studies on this point are discrepant. Thus Gaw's study of neglected canal boat children (who attended school only 5 per cent of the time) in England showed an *IQ* of 82 in the performance tests as against an *IQ* of only 69 in the Binet Here the impoverished environment affected the Binet rating more than the motor rating in contrast with the Iowa finding

F Conclusions

This presentation of findings will be concluded with two brief comments, the one largely theoretical in character and the other practical in its implications

1 What is involved in the concept of fluctuations or variations of intelligence? What do the fluctuations postulate? What is the nature of this abstraction that we denominate intelligence that is subject to fluctuations? Is the underlying "thing-in-itself," the basic substrate or energy that expresses itself in intellectual functioning, a fixed, static quantum, or is it a constantly changing force? Is the underlying strength of intelligence uniform or is it fluctuating?

Let us say forthright that the only phase of intelligence open to test exploration is that which functions in test situations or which can be observed in overt behavioral reactions. We may refer to this as the phenomenal and knowable aspect of intelligence and personality as opposed to the noumenal self which, whatever it may be, is inaccessible to the examiner Therefore, any conclusions reached must necessarily relate to the phenomenal or functional aspect of intelligence. To venture upon conclusions regarding the ultimate nature of the vital energy at the root of overt intellectual activity would lead us into a morass of speculation that we will do well to avoid. There is no way of experimentally exploring pure intelligence in a vacuum nor is there any way for intelligence to develop or manifest itself in a vacuum We can only study intelligence as it functions in behavior To assert that fluctuations in intelligence as evidenced in test findings are paralleled by corresponding fluctuations in the underlying reality to which we popularly

apply the term intelligence, whatever it may be, may be true, but it would constitute a speculative leap in the dark to so assert categorically and would raise questions that had better be reserved for the metaphysician After all, the psychological examiner is able to measure only the psychophysical reactions of the organism and his conclusions should be limited to observable phenomena It is conceivable that the reactions or responses of the organism might be greatly interfered with or might be greatly facilitated by organic and extraorganic conditions without producing fundamental modifications in the underlying intellectual energy To conclude from test findings, even though they have been adequately documented and are not spurious because of the imperfections in the tests or in the rating techniques or for other reasons, that an individual's intellect has improved 50 *IQ* points or deteriorated 50 *IQ* points may represent an inductive leap beyond the premises All that the deterioration revealed by the tests may mean is that the outlets of expression have been impaired or blocked by obstructions, temporary or otherwise *IQ* gains might likewise be due to facilitations or excitations of the neuromuscular apparatus, temporary or otherwise.

All this aside, however, it is indisputable that fluctuations do occur in tested intelligence, whether the children are normal or abnormal, whether the retests are given by the same or different examiners, or whether the retests are given infrequently or in a long series of retests at six months' intervals. That the variations are sometimes so large as to lead to highly discrepant diagnoses, depending upon the test result that is accepted, is equally notorious That the interpretation of experimentally established fluctuations in *IQ*'s is not readily apparent, is equally patent The fluctuations may be genuine or illusory They may be partly correct and partly misleading They may be exaggerated or minimized by the testing or scoring techniques They may be caused by endogenous or exogenous factors They may represent the effects of practice or of emotional inhibitions They may be permanent or temporary They may be subject to remediation or they may prove to be highly resistant to modification Broad diagnostic and prognostic generalizations should, therefore, not be ventured in clinical work without a careful consideration of all the factors, external and internal, that affect responses to psychological tests of whatever nature In other words, psychological diagnosis and prognosis should be based upon the

thorough-going application of case-study procedures. In such procedures the use of standardized psychological tests, whatever their limitations, is of basic importance

2. The practical bearing of all this is that inferences from test findings regarding a child's intellectual ability should not be hastily reached and momentous decisions based on such findings should be withheld in the absence of sufficient corroborative evidence One must not confide too implicitly in test scores or place implicit dependence on *IQ* predictions Although the *IQ* in general may be fairly reliable, and although the *IQ* is perhaps the most valuable single measure we possess of relative intelligence, after all the *IQ* is only one factor in mental diagnosis In a sense, an *IQ* is merely an *X* to be interpreted in the light of all the synoptic findings bearing on the case Viewed in isolation, it may become little better than a cryptic symbol and a source of confusion. We have found reasons to maintain that *IQ*'s are sometimes spuriously low or spuriously high without full realization of the fact by the examiner Let us, therefore, emphasize that classifications and assignments should be tentative, and subject to revision in the light of a fuller knowledge of the examinee.

That the IQ may give a correct measure of the child at the time, no matter what the *IQ* was a year earlier or will be next year, is no doubt true in the majority of cases But spurious *IQ*'s do not give a correct appraisal at the time of the test

The best way to avoid errors inherent in the scale of tests is to secure corroborative evidence from other tests and from a thorough case study of the individual. But even when the best diagnostic devices have been used a reasonable amount of caution should always be observed in making psychological diagnoses and prognoses of children.

It is particularly important that the classroom teacher should be freed from doctrinaire and fallacious ideas regarding the inerrancy of *IQ*'s as well as of all other kinds of *Q*'s Always the teacher and the educator should give the child the benefit of the doubt. An *IQ* should never serve as an excuse for complacency or justification for inaction. Ample opportunities should be afforded young retardates and deficients in the literary subjects, using the best remedial techniques, concrete procedures, and interest-provoking activity programs, before such instruction is abandoned and the emphasis diverted to motor training

Doubtless, many young children have been institutionalized as feeble-minded and thereafter retained permanently in institutions for such children who would have improved to such an extent in a superior public school environment that they would not have been classified as feeble-minded after several years of efficient instruction On the other hand, many young children have been excluded from public school special classes on the basis of an arbitrary *IQ* standard, usually a Binet *IQ* of 50 and below, who would have made reasonable progress in such classes. The writer has known many cases in point with *IQ*'s below 50 who have made greater progress than others with *IQ*'s in the fifties or sixties, either because they possessed counter-balancing personality traits of greater value than the low *IQ* or because the *IQ* was spuriously low as proved by the children's subsequent development.

The writer has for many years felt that too many children are being sent to residential institutions for the feeble-minded where they may receive inadequate instruction because of the crowded conditions or because of substandard instruction and where the institutional life tends to make against the development of initiative, resourcefulness, or independence If the Iowa results mean anything, even after you have discounted the possible exaggerations, they mean this, that young children, whether they test normal or subnormal, should not be regimented into a form of congregate existence that is barren of cultural and intellectual stimulation It is no doubt true that many institutions afford not only better shelter but also better social and intellectual stimulation than many impoverished and incompetent homes But the truth remains, nevertheless, that a good home is superior to a good institution in the intellectual stimulation and personality development of subnormal as well as normal children, at least if the children are above the level of feeble-mindedness The first line of attack in the education of subnormals is the public schools rather than the colonies. Even children with *IQ*'s below 50 should not be arbitrarily denied the advantages of the special classes without an adequate period of probation. Not infrequently the *IQ*'s are fallaciously low, as has been emphasized Without adequate environmental stimulation potentialities have no chance to develop Therefore, the child is entitled to a reasonable trial before irrevocable decisions are reached concerning his future Residential institutions play a very important rôle in the program for

salvaging mentally defective children, but they should not usurp the function of the public school nor should the public schools force children into these institutions so that they may justify themselves in shirking their responsibilities toward "all the children of all the people." Public school officials exist who are perfectly willing to be relieved of the responsibility for educating defectives What, then, should be the prime function of the residential training schools so far as concerns the education of mentally defective children? The institutions should play the rôle of auxiliary training schools for idiots and low imbeciles and for mentally defective children who are already anti-social or delinquently or criminalistically inclined, who are unprotected or unsupported or who make too great demands on an overworked mother or who are victims of educational neglect because the public schools cannot or will not cater to their needs

One final reflection on the writer's credo, based on experience with thousands of defectives, concerning the effects of environmental stimulation: Environmental stimulation is of paramount importance in bringing intellectual potentialities to fruition. Without environmental stimulation intellectual atrophy would inevitably result, intelligence cannot flourish in a vacuum Nevertheless, neither the physical nor cultural environment is omnipotent Native endowment sets insuperable limits so far as concerns the development of intelligence. Constitutional limits exist that not even a superior environment can completely overcome. Therefore, the impression should not be allowed to gain currency, that a superior environment will transform all mediocre children into geniuses, or that a barren environment will reduce all such children to the state of amentia While the transformation wrought by education often looks like the creation of new powers, education perhaps only facilitates the unfoldment of existent powers In the complete absence of potentialities to be released and unfolded, education is as impotent as the fruition of potentialities in vacuo Nevertheless, after conceding all this, we must recognize that the limits set by heredity are not as rigidly fixed as we once believed. Without a favorable social and educational environment, many who have attained eminence would have remained mediocre. Education is a constructive force no less for potential morons than for potential geniuses.

References

1. Pintner, R Intelligence Testing New York Holt, 1923
2. Skeels, H M Mental development of children in foster homes *J Consult Psychol*, 1938, **1**, 33-43
3. Skeels, H M, Updegraff, R, Wellman, B, & Williams, H M A Study of Environmental Stimulation *Univ Iowa Stud*, 1938, **15**, No 4
4. Wallin, J. E W Problems of Subnormality. New York World Book, 1917.
5. ——— Wide range versus narrow range Binet-Simon testing *J Delin*, 1917, **2**, 315-330
6. ——— Clinical and Abnormal Psychology. Boston Houghton-Mifflin, 1927
7. Wellman, Beth L Mental Growth from Preschool to College *J Exper Educ*, 1937, **6**, 127-138
8. ——— Our changing concept of intelligence *J Consult Psychol*, 1938, 1, 97-107

Department of Special Education
1406 Washington Street
Wilmington, Delaware

The Journal of Genetic Psychology, 1940, 57, 393-414

ADOLESCENTS' DISLIKES REGARDING PARENTAL BEHAVIOR, AND THEIR SIGNIFICANCE*[1]

Nebraska Agricultural Experiment Station, University of Nebraska

LELAND H. STOTT

Parents generally are coming to realize the importance of the person-to-person relationships and the general social environment of the home in relation to their children's personal development. Parent-child relationships are claiming the attention of greater numbers of investigators, and with the wider dissemination of research information, child rearing is more generally being regarded as a job which requires some study.

At the same time, parents perhaps more often than they realize, through their own attitudes, personal habits and social conduct are determining not only their children's attitudes toward them, but also their reactions and adjustments to life in general. Studies have shown, among other things, that young children most frequently like the parent best who gives them most satisfaction and contributes most to their physical comfort, who plays with them and expresses most affection for them and who punishes them least (4, 5). It has also been shown that incompatibility between parents themselves, as well as the lack of comradeship between adolescents and their parents, was associated with the rebellious attitude against parental discipline (8) and with delinquency (6) on the part of the youngsters. Other findings indicate, further, that children tend to adopt the behavior patterns and moral standards of their parents, and thus to become more like their parents in regard to those patterns and standards (3). These findings suggest that even though children may dislike and resent certain modes of behavior and certain attitudes in their parents, they may and perhaps often do, adopt those same modes

*Received in the Editorial Office on September 18, 1939.

[1]Published with the approval of the Director as Paper No. 248 Journal Series, Nebraska Agricultural Experiment Station. This report is one of a series dealing with the results of a research project, the general subject of which was the relation between home-environmental influences and personality adjustments in children.

of behavior and those same attitudes as their own in spite of their conscious dislike for them as they observe them in their parents

What young people dislike in the behavior of their parents, as well as what they admire, then, probably are matters of some importance in relation to their own personal development. Just how important this relationship is, and whether or to what extent it varies from one general home situation to another, or with the character of the disliked behavior or trait, are questions with which the present study deals

An attempt was made to inventory and classify the items of parental behavior which, under certain specified conditions, adolescent boys and girls of three different cultural home settings said they did not like in their own parents, and to study the relative importance of the different classes of items in relation to the personality adjustments of the youngsters The data were the responses of the subjects to the questionnaire items, *"What does your mother do that you do not like?"* and *"What does your father do that you do not like?"* together with their scores on scales designed to measure several different aspects of personality [2]

In the first section of this report are presented the relative frequencies of criticisms of mother and of father by children from farm, small town, and city homes, and by the two sexes The question of the relation between criticizing or not criticizing the parents and development of the subjects is next considered A qualitative classification of the specific items of behavior criticized by the boys and the girls of the three cultural groups is then presented and certain group differences in the nature of the more frequently criticized behavior items are pointed out. These differences are considered in relation to the differences among the cultural groups in degree of observed correlation between offering criticisms of parents and unfavorable showing on the personality scales

The Frequency of Criticism

Following each of the two questionnaire items with which we are here concerned, was provided a line on which the subject might write his criticism. No check list was provided, hence nothing was

[2]Descriptions of the home-life questionnaire, the tests and scales used and the general procedures followed are given elsewhere (7)

suggested to the subject. He was free to leave the space blank, to write the word "nothing" or to mention anything about the parent which might occur to him, with which he was not pleased. Of the 1,878 subjects who filled in the questionnaire, 675 (35 9 per cent) mentioned something in mother's behavior not liked and 679 (36 2 per cent) mentioned something in father's behavior not liked Nearly two-thirds of the subjects, in other words, either left the space blank or stated that they had no criticisms to offer

Of this total group of young people, 694 were living with their parents on farms, 639 were living in small towns, and 545 were living in the city of Omaha The tendency was for the small town youngsters more frequently to criticize their parents The percentages of the residence groups that offered criticisms of mothers behavior were 39 0 for the town, 35 6 for the farm, and 32 8 for the city The difference in proportion between town and city groups is fairly reliable statistically (more than two times its standard error). Thirty-nine per cent of the town subjects, 37 0 per cent of the farm subjects and 31 7 per cent of the city subjects criticized their father's behavior The difference between the small town and the city groups is again reliable

When the sex groups are compared it is seen that the girls were responsible for most of the differences between the residence groups The small town girls saw more to criticize in their parents than any other grouping of subjects Of this group, 46 3 per cent, as compared with 35 5 per cent of the farm girls, and 35 6 of the city girls, mentioned things not liked in their mothers As to frequency of criticism of father, the town and farm girls were more nearly alike, but the city girls were significantly different Only 21 2 per cent of this group, as compared with 44 2 and 40 2 per cent of the town and farm girls, offered criticisms of their fathers When all of the 956 girls were grouped together and compared with the 922 boys, the girls were shown to be significantly more disposed to mention something not liked in both parents than were the boys The percentages were 34 3 and 32 4 for criticism of mother and 39 2 and 33 0 for criticism of father In Table 1 are shown the percentages of all the groupings of the subjects who offered criticisms Table 2 gives the significant group differences

TABLE 1

PERCENTAGES OF THE VARIOUS GROUPINGS OF SUBJECTS THAT CRITICIZED MOTHER'S AND FATHER'S BEHAVIOR

Grouping	N	Per cent criticized mother	Per cent criticized father
Farm boys	328	35.7	33.5
Farm girls	366	35.5	40.2
Total farm	694	35.6	37.0
Town boys	302	30.8	33.1
Town girls	337	46.3	44.2
Total town	639	39.0	39.0
City boys	292	30.5	32.2
City girls	253	35.6	21.2
Total city	545	32.8	31.7
Total boys	922	32.4	33.0
Total girls	956	39.3	39.2
Total subjects	1878	35.9	36.2

TABLE 2

SIGNIFICANT GROUP DIFFERENCES IN PER CENT CRITICIZED MOTHER AND FATHER

Significant group differences in per cent criticized mother			Significant group differences in per cent criticized father		
Groups compared	*Diff*	$\frac{Diff.}{\sigma_{diff}}$	Groups compared	*Diff*	$\frac{Diff}{\sigma_{diff}}$
Town girls>town boys	15.5	4.08	Town girls>town boys	11.1	2.90
Town girls>farm girls	10.8	2.93	City boys>city girls	11.0	2.93
Town girls>city girls	10.7	2.64	Farm girls>city girls	18.0	4.96
Total town>total city	6.2	2.22	Town girls>city girls	23.0	6.17
Total girls>total boys	6.9	3.12	Total town>total city	7.3	2.64
			Total girls>total boys	6.2	2.81

RELATION BETWEEN CRITICIZING PARENTS AND PERSONALITY ADJUSTMENTS

The correlation analyses revealed that the relationship between mentioning something not liked in parents and the personality variables as measured in the subjects were small and mostly insignificant. The three cultural (residence) groups, however, differed considerably one from another in regard to the particular variables which

gave reliably large correlation The sex groups also varied in that regard. In Table 3 are shown all of the significant[3] correlations found for all subject groupings

TABLE 3

SIGNIFICANT CORRELATIONS BETWEEN MENTIONING SOMETHING NOT LIKED IN PARENTS' BEHAVIOR AND SCORES ON THE PERSONALITY SCALES

Variables	Farm Boys	Farm Girls	Town Boys	Town Girls	City Boys	City Girls
Criticisms of Mothers' Behavior						
Personal adjustment (Maller "A" score)					— 35	
Ethical judgment (Maller "E" score)					— 28	— 22
Independence					— 25	
Personal responsibility					— 26	
Attitude toward work			— 22		— 20	
Appreciation of home life		— 21		— 23	— 26	— 34
Average personality rating[1]			— 24		— 35	— 24
Criticisms of Fathers' Behavior						
Personal adjustment (Maller "A" score)					— 24	
Ethical judgment (Maller "E" score)					— 21	
Attitude toward work			— 20			
Appreciation of home life	— 25			— 20	— 24	— 35

[1]The "average personality rating" was the average of the subject's standard scores (modified) on all the personality scales which he completed satisfactorily

The greatest number of significant correlations appeared for the city subjects For this group there were 14 correlations of — 20 or greater as compared with only five and two respectively for the small town and farm groups When the sex groups were compared as to the number of significant correlations, it was the boys that stood out Of the total of 21 correlations greater than — 20, 14 were contributed by the boys, 10 of which came from the city boys and only one from the farm boys

[3]For purposes of this study no correlation smaller than 20 was regarded as significant In some instances correlations of lesser magnitude were "statistically significant" according to Fisher's criterion of significance (2, pp 335-336 and 548) but they are not included here Every correlation included in Table 3 is more than twice its standard error

The analysis also suggested that to have something to criticize in the mother was, in general, of greater significance in its relationship to low personality scores than to have something to criticize in father Fourteen of the total of 21 correlations greater than — 20 involved criticisms of mother Seven of these 14 came from the city boys, while not a single one came from the farm boys This constitutes, perhaps, the most striking contrast of any The mention of something not liked in mother by the farm boys was of no significance in relation to any of the personality scores but with the city boys such a response was significantly associated with low scores in (*a*) personal adjustment (Maller), (*b*) ethical judgment (Maller), (*c*) independence (self reliance) in meeting personal difficulties, (*d*) personal responsibility in maintaining satisfactory personal relationships with others, (*e*) attitude toward work, (*f*) appreciation of (attitude toward) home life and (*g*) average personality rating

There is, of course, no way of determining directly whether, or to what extent these correlations represent causal relationships It seems reasonable, however, to suspect that there is something characteristic, for example, in the city home situation, perhaps in the prevailing nature of the parent-child relationships, or in the social or personal behavior patterns commonly adopted by city mothers, which actually renders the mothers' conduct a more serious matter in the lives and the personal adjustments of their sons than is the case in the farm home situation. It was deemed worthwhile therefore, to make a comparative study of the items of parental behavior more frequently criticized by boys and by girls of the three cultural groups in the hope that it might throw some light upon this problem

Group Comparisons as to Frequency of Mention of the Various Items of Parental Behavior

An attempt was made to classify qualitatively the things criticized in parents The criticizing seemed roughly to fit into the following groups. (*a*) those having to do with discipline and control and other personal relationships with the child himself, (*b*) those regarding temperamental behavior and traits affecting members of the family and others; (*c*) those regarding the personal habits and conduct of

the parent; (*d*) those concerning the parent's health, and emotional adjustment and control, (*e*) those regarding the parent's ideas, attitudes or beliefs, (*f*) those regarding work and self-sacrifice; (*g*) those regarding the social adjustments and activities of the parent, (*h*) a miscellaneous group, mostly trivial All of the criticisms mentioned of mothers' behavior and of fathers' behavior are listed according to this classification together with their respective frequencies in Tables 4 and 5 This, of course, is merely one of a number of

TABLE 4

INVENTORY OF ITEMS OF MOTHERS' BEHAVIOR REPORTED AS NOT LIKED BY ADOLESCENTS OF FARM, SMALL TOWN AND CITY HOMES

Items reported	Frequency of report					
	Farm		Town		City	
	Boys	Girls	Boys	Girls	Boys	Girls
I Discipline and control, and other personal relationships between parent and child						
Scolds	15	18	4	22	3	8
Threatens	1					1
Restricts, or refuses to grant privileges	3	1	2	2	1	
Corrects me		5			1	
Strict	4	4	5	4		3
Too easy					1	
Makes me work	5	3	7	9	4	
Makes me stay home	3	6				
Makes me go to bed or get up (early)	2	1	4	1		
Makes me go to school	3					
Makes me hurry		1				2
Leaves me home	2					
Doesn't allow me to "date"						1
Doesn't let me go out often				9		
Objects to late hours				4		
Doesn't want me to work		1				
Tries to assume my duties			3			
Tries to push me ahead					2	
Doesn't give me enough money				2		
Has me do embarrassing things					1	
Forces me to do something				2		4
Tells me what to wear						1
Expects too much of me				1		
Asks personal questions	1	3				3
Borrows my money					1	

TABLE 4 (*continued*)

Items reported	Farm Boys	Farm Girls	Town Boys	Town Girls	City Boys	City Girls
Refuses to let me bring my friends						1
Total frequencies of class I	39	43	25	56	14	24
Per cent of total criticisms	33 3	33 1	26 9	35 9	15 7	26 7
II Temperamental traits and behavior affecting members of family and others						
Complains, nags, crabs	2	13	2	14	4	5
Finds fault, criticizes				7		2
Blames wrong person			2	2		1
Argues	1			1		1
Teases				1		
Quarrels			1			
Fights					1	
Wants last word					1	
Inconsiderate	2		4	3		2
Bossy	1		1	3	2	
Partial	3	5	1	6	4	1
Cross	1					
Uncooperative						1
Stubborn				2		
Total frequencies for class II	10	18	11	39	12	13
Per cent of total criticisms	8 5	13.8	11 8	25	13 5	14 4
III. Personal habits and conduct						
Smokes					5	5
Drinks						1
Swears	1				2	
Talks too much	6	2	5	1	8	
Talks too much on telephone						5
Gossips				2		1
Brags, exaggerates			1		2	1
Mannerisms					4	
Neglects herself		1		1		
Repeats things						1
Impatient				2		
Extravagant			1			
Noisy				2		
Forgetful			1			1
Total frequencies for class III	7	3	8	8	21	15
Per cent of total criticisms	6	2 3	8 6	5 1	23 6	16 7
IV Health, emotional adjustment and control						
Gets sick					1	
Nervous	1	5	2	7	4	5
Worries	6	8	3	3	4	2

TABLE 4 (*continued*)

Items reported	Frequency of report Farm Boys	Farm Girls	Town Boys	Town Girls	City Boys	City Girls
Loses temper	8	7	4	5	5	5
Jealous		1				
Feelings easily hurt		1				
Dissatisfied					1	
Total frequencies for class IV	15	22	9	15	15	12
Per cent of total criticisms	12.8	16.9	9.7	9.6	16.9	13.3
V. Ideas, attitudes, beliefs						
Old, funny ideas	4	1	3	3		
Gets wrong ideas		1				
Draws hasty conclusions		1				
Too conservative			1			
Takes things too seriously				1		
Doesn't understand my friends				1		
Total frequencies for class V	4	3	4	5	0	0
Per cent of total criticisms	3.4	2.3	4.3	3.2	0	0
VI. Work, self sacrifice						
Works too hard	9	10	9	4	9	3
Works outside home	4		3	9		1
Works when sick						1
Deprives self of necessities					2	
Runs home inefficiently						1
Meals irregular, not on time	1					1
Not as sanitary as should be		1				
Total frequencies for class VI	14	11	12	13	11	7
Per cent of total criticisms	12	8.5	12.9	8.3	12.4	7.8
VII. Social adjustments						
No social life	4		5			
Stays home too much		2		2	2	
Doesn't mingle with friends						1
Too much social life					3	
Has undesirable friends						1
Total frequencies for class VII	4	2	5	2	5	2
Per cent of total criticisms	3.4	1.5	5.4	1.3	5.6	2.2
VIII. Miscellaneous						
Trivial things	24	28	19	18	11	17
Per cent of total criticisms	20.5	21.5	20.4	11.5	12.4	18.9
Total criticisms of mothers' behavior	117	130	93	156	89	90

possible groupings, and possibly not the best one for our purpose. The classes, it will also be noted, are somewhat over-lapping, and in a number of instances an item was arbitrarily placed in one class when it might logically have belonged just as well in another.

As was shown in Tables 1, 2, and 3, the residence groups and the

TABLE 5

INVENTORY OF ITEMS OF FATHERS' BEHAVIOR REPORTED AS NOT LIKED BY ADOLESCENTS OF FARM, SMALL TOWN AND CITY HOMES

	Frequency of report					
	Farm		Town		City	
Items reported	Boys	Girls	Boys	Girls	Boys	Girls
I Discipline and control, and other personal relationships between parent and child						
Scolds, "lectures"	12	12	5	15	9	10
Threatens						
Restricts privileges	2	3	5	5	8	
Corrects me					2	
Strict	5	4	4	6		4
Too easy		2				
Makes me work	4	3	1	1	1	
Won't let me work		1				
Leaves me home	2					
Makes me stay home	3	7				
Makes me get up early	2					
Won't let me "go out", "date"				7		1
Teases, kids me				3		
Picks on me	5					
Doesn't care about me going to high school		2				
Doesn't give me money			3			
Expects too much of me					2	
Forces me to do something				2		
Protects me too much				4		
Total frequencies for class I	35	34	18	43	20	17
Per cent of total criticisms	31.8	23.1	18	28.9	21.3	21.5
II Temperamental traits and behavior affecting members of the family and others						
Argues		1		1		1
Grouchy, cross, finds fault	6		4		6	
Nags, crabs, finds fault, criticizes		2		5		6
Cross		5		3		
Bossy	1	1	2	2	2	
Inconsiderate		8	3	7	1	5
Cruel			1			
Partial	1		1	4		
Rude		2				
Stubborn	1	1	1	1		2
Sarcastic					1	2
Teases		3				5
Quarrels			1			
Fights					1	
Plays jokes			1			
Isn't companionable	1					

TABLE 5 (*continued*)

Items reported	Frequency of report Farm Boys	Farm Girls	Town Boys	Town Girls	City Boys	City Girls
Makes too many promises			1			
Insistent				3		
Expects too much of mother				1		
Selfish, stingy				2		1
Total frequencies for class II	10	23	15	29	11	22
Per cent of total criticisms	9.1	15.6	15	19.5	11.7	27.8
III Personal habits and behavior						
Drinks	1	1	3	2	9	1
Smokes and (or) chews	14	25	23	14	7	6
Swears, curses	3	3		4	2	3
Brags, boasts, exaggerates			1	1	2	3
Gossips						1
Talks too much	1	3	2	2	5	
Noisy				3		
"Hollers"					2	
Gets home late	1		1			
Not home enough		9				
Bad table manners		2				
Mannerisms			3	1	2	
Forgetful	1		1			
Too particular						2
Puts things off						1
Dirty	1					
Extravagant, borrows					3	
Way he drives car, too fast		2		2		1
Total frequencies for class III	22	45	34	29	32	18
Per cent of total criticisms	20	30.6	34	19.5	34	22.8
IV Emotional adjustment and control						
Loses temper	13	17	7	7	11	3
Worries	1	3		3		
Nervous						2
Jealous		1				
Feelings easily hurt		1				
Moody		1				
Total frequencies for class IV	14	23	7	10	11	5
Per cent of total criticisms	12.7	15.6	7	6.7	11.7	6.3
V Ideas, attitudes, beliefs						
Old fashioned ideas			2	3		
Radical			1			
Gets wrong ideas		1				
Takes things too seriously				1		
Total frequencies for class V	0	1	3	4	0	0
Per cent of total criticisms	0	0.7	3	2.7	0	0

TABLE 5 (*continued*)

Items reported	Frequency of report: Farm Boys	Farm Girls	Town Boys	Town Girls	City Boys	City Girls
VI Works, self-sacrifice						
Works too hard	14	8	4		3	2
Lazy	1			2	2	
Selfish			2			
Not as sanitary as should be		1				
Total frequencies of class VI	15	9	6	2	5	2
Per cent of total criticisms	13.6	6.1	6	1.3	5.3	2.5
VII Social matters						
No social life	2					
Stays home too much		2		4		
Unsocial						2
Has undesirable friends						1
Total frequencies for class VII	2	2	0	4	0	3
Per cent of total criticisms	1.8	1.4	0	2.7	0	3.8
VIII Miscellaneous						
Trivial things	11	10	10	21	13	12
Occupation			2	5		
Regresses				2		
Way he drives car			3			
Doesn't live with us			1			
Not home enough, travels	1		1		2	
Total frequencies for class VIII	12	10	17	28	15	12
Per cent of total criticisms	10.9	6.8	17	18.8	16	15.2
Total criticisms of fathers' behavior	110	147	100	149	94	79

sex groups differed, not only in regard to frequency of reporting things not liked in parents, but also in regard to the degree to which criticizing parents was correlated with inadequate personality adjustment. Although smaller percentages of the city subjects than those of farm or town reported criticisms, most of the correlations shown in Table 3 were from the city subjects. The boys, when taken together, likewise offered fewer criticisms than the girls, yet criticism of parents was more often found to be correlated with low personality scores in boys than in girls. The percentages of all subjects who offered criticisms of mother and father were not significantly different, but criticizing mother was reliably related to more of the personality variables in the subjects than was criticizing father. Our present task, then, is to compare the groups as to frequency of criticism in each *qualitative class* and thus to search for

group differences which might have some bearing on the problem of the significance of parental behavior in the adjustments of normal adolescents.

1. *Criticisms of Mothers and Fathers Compared*

In Table 6 are shown the percentages for each class of the total

TABLE 6

COMPARISON OF RELATIVE FREQUENCIES OF CRITICISM OF EACH QUALITATIVE CLASS OFFERED AGAINST THE MOTHERS AND FATHERS BY ALL SUBJECTS COMBINED

Total criticisms of mother, 675 Total criticisms of father, 679

Criticized parent	Per cent of total criticisms of parent	*Diff*	*Diff* / σ_{diff}
	I. *Matters related to discipline and control*		
Mother	29.8	5.2	2.16
Father	24.6		
	II. *Temperamental traits and behavior*		
Mother	15.3	−0.9	
Father	16.2		
	III. *Personal habits and conduct*		
Mother	9.2	−17.3	−8.52
Father	26.5		
	IV. *Emotional adjustment and control*		
Mother	13.0	2.7	
Father	10.3		
	V. *Ideas, attitudes, beliefs*		
Mother	2.4	1.2	
Father	1.2		
	VI. *Work, self-sacrifice*		
Mother	10.1	4.4	3.01
Father	5.7		
	VII. *Social adjustments and behavior*		
Mother	3.0	1.4	
Father	1.6		
	VIII. *Miscellaneous, trivial*		
Mother	17.3	3.5	
Father	13.8		

criticisms offered against the mothers and against the fathers. Differences in percentage and critical ratios for the reliable differences are also shown. The most striking difference was found in percentages for Class III. Fathers, almost three times as frequently as mothers, were criticized because of some personal habit

or item of personal conduct Only two other percentage differences were statistically reliable A somewhat greater proportion of the complaints against mothers had to do with discipline and methods of control of the children This is not surprising since, according to the results of a questionnaire item asking about punishment in the home, 64 per cent of all the home punishment of both boys and girls was administered by the mother as compared with only 31 per cent administered by the father.[1] The girls were punished four times as often by the mother as by the father, according to their reports

The other significant difference was in the relative frequency of complaint regarding work habits, and tendencies of overwork and self-sacrifice in parents (Class VI) The percentages of total criticisms which were included in this class for both parents, however, were quite small, only 10 1 and 5 7 for mothers and fathers respectively. These comparisons are interesting but they, in themselves, suggest little by way of an answer to the question of why criticizing mother was related to more of the personality scores than criticizing father.

2. *Total Sex Groups Compared as to Criticism Offered*

As was shown in Tables 1 and 2, the girls, when taken together from all three residence groups, were more prone to criticize both parents than were the boys. Significant relationships between criticizing parents and the personality variables, however, were more numerous with the boys than with the girls (Table 3) In Table 7 are shown comparisons of percentages of total criticisms from boys and from girls which belonged to the various classes of criticism These are given separately for mothers and fathers. Only four significant percentage differences appear in the whole table Reliably greater proportions of the girls' criticisms than of the boys' criticisms of both parents had to do with temperamental traits and difficulties (Class II). The tendency (differences not reliable) was also for the girls more frequently than the boys to report items of Class I, matters of discipline and control The common element

[1]Bowers (1) reported that according to the answers of his subjects, 54 per cent of their punishment was administered by their mothers as compared with 31 per cent by their fathers

TABLE 7

COMPARISON OF THE SEX GROUPS AS TO RELATIVE FREQUENCY OF CRITICISMS OF EACH CLASS OFFERED AGAINST MOTHERS AND AGAINST FATHERS

Criticized Mother, boys 299, girls 376. Criticized Father, boys 304, girls 375

Sex Group	Criticisms of mother: Per cent of total criticisms	*Diff*	*Diff* / σ_{diff}	Criticisms of father: Per cent of total criticisms	*Diff*	*Diff* / σ_{diff}
	I *Matters related to discipline and control*					
Boys	26.1	—6.6		24.0	—1.1	
Girls	32.7			25.1		
	II *Temperamental traits and behavior*					
Boys	11.0	—7.6	—2.84	11.8	—7.9	—2.86
Girls	18.6			19.7		
	III *Personal habits and conduct*					
Boys	12.0	5.1	2.23	28.9	4.4	
Girls	6.9			24.5		
	IV *Emotional adjustment and control*					
Boys	13.0	0.0		10.5	0.0	
Girls	13.0			10.1		
	V *Ideas, attitudes, beliefs*					
Boys	2.7	0.6		1.0	—0.3	
Girls	2.1			1.3		
	VI *Work, self-sacrifice*					
Boys	12.4	4.2		8.6	5.1	2.73
Girls	8.2			3.5		
	VII *Social adjustments and behavior*					
Boys	4.7	3.1		0.6	—1.8	
Girls	1.6			2.4		
	VIII *Miscellaneous, trivial*					
Boys	18.1	1.3		14.5	1.8	
Girls	16.8			13.3		

in the two classes is of course that both involved matters which affected the person-to-person or social relationships between the criticized parent and the child or other members of the family. The girls, definitely more frequently than the boys, objected to such things in the behavior of their parents.

On the other hand, relatively greater proportions of the boys than of the girls who offered criticisms of their parents, mentioned matters pertaining to the parent's personal habits and conduct. The difference was significantly large for criticisms of mother. This suggests that it is more frequently important to the adolescent boy than to the adolescent girl that the mother's habits and personal conduct be decorous, and such that he can be proud of her.

To sum up regarding the comparisons between the total boy and

the total girl groups, then, we may say that the adolescent boy, although not so prone to report things not liked in his parents, tends more than the girl to report items which are more truly criticisms of the parent himself or herself, and not nearly so frequently as the girl does he report items pertaining to the personal and social relationships of the parent with the other family members That the reported criticisms of boys are to them more serious on the average, is suggested by the greater number of significant correlations between criticism and personality adjustments as measured in this study.

3 *The Cultural Groups*

In Table 8 may be found the percentages which fell in each qualitative class of the total criticisms of parents given by each cultural group Of the total criticisms reported by the city group only 21 3 per cent fell in Class I as compared with 28 5 per cent and 30 3 per cent for the town and farm groups As to proportion of criticisms falling in Classes II and III, the farm group was lowest In the case of Class III the proportion for the city group (24 4 per cent) was significantly larger than those of the other two residence groups. As to Class IV the percentages were 14 9 for the farm subjects, 8.2 for the town subjects, and 12 2 for the city subjects The difference in percentage between the farm and city groups was not significant.

In connection with criticism Classes I, II, and III then, were found the most significant differences between the farm and the city subjects, the two residence groups which showed greatest contrast in the number of significant correlations (Table 3)

Classes II and III accounted for 40.9 per cent of the total criticisms offered by the city subjects and only 27 4 per cent of those offered by the farm subjects, while Class I accounted for only 21 5 per cent of the total criticisms of the city subjects as compared with 30 3 per cent of those of the farm subjects These figures indicate then that those city youngsters who reported things objected to in their parents' behavior, tended definitely to report traits and difficulties of a temperamental nature and personal habits and items of conduct which they regarded as objectionable, reliably more frequently than did the farm youngsters The latter group, on the

TABLE 8

COMPARISON OF THE RESIDENCE GROUPS AS TO RELATIVE FREQUENCY OF CRITICISM OF EACH QUALITATIVE CLASS

Total criticisms, farm 504, town 498, city 352

Residence group	Per cent of total criticisms from group	Diff		$\frac{Diff}{\sigma_{diff}}$
	I. *Matters related to discipline and control*			
Farm	30.0	F>T	1.8	
Town	28.5	F>C	9.0	3.02
City	21.3	T>C	7.2	2.50
	II. *Temperamental traits and behavior*			
Farm	12.1	T>F	6.8	2.98
Town	18.9	C>F	4.4	
City	16.5	T>C	2.4	
	III. *Personal habits and conduct*			
Farm	15.3	T>F	0.6	
Town	15.9	C>F	9.1	3.26
City	24.4	C>T	8.5	3.02
	IV. *Emotional adjustment and control*			
Farm	14.9	F>T	6.7	3.33
Town	8.2	F>C	2.7	
City	12.2	C>T	4.0	
	V. *Ideas, attitudes, beliefs*			
Farm	1.6	T>F	1.6	
Town	3.2	F>C	1.6	
City	0.0	T>C	3.2	4.10
	VI. *Work, self-sacrifice*			
Farm	9.7	F>T	3.1	
Town	6.6	F>C	2.6	
City	7.1	C>T	0.5	
	VII. *Social adjustments and behavior*			
Farm	2.0	T>F	0.2	
Town	2.2	C>F	0.8	
City	2.8	C>T	0.6	
	VIII. *Miscellaneous, trivial*			
Farm	14.7	T>F	1.8	
Town	16.5	C>F	0.9	
City	15.6	T>C	0.9	

other hand, tended more frequently to report matters which often were not criticisms of parental behavior at all, but were rather superficial objections to parents' methods of discipline and control or to other personal matters, often trivial, between the subject and his parent, such, for example, as "asks personal questions," "tries to push me ahead," or "tries to assume my duties." Our city sub-

jects, then, offered fewer criticisms of their parents behavior as compared with those from farm homes, but the criticisms which they reported tended more frequently to be concerned with matters of a more serious nature—matters which presumably were sources of difficulty to them in their own personal adjustments The criticisms of the farm subjects, on the other hand, were of a less serious nature and hence not related to their own personal adjustments

4 *City Boys and Farm Boys Compared in Regard to Criticisms of Mother*

The two groups which showed the greatest contrast in regard to the number and size of the significant correlations between criticizing parents and personality scores (Table 3) were the city boys and the farm boys. This contrast was particularly striking in regard to their criticisms of mother In Table 9 are shown the comparisons of these two groups as to relative frequencies of complaints of the various classes against their mothers. Because of the small numbers (117 farm boys and 89 city boys) only two of the percentage differences were statistically significant Certain interesting comparisons, however, may be noted in the table. Items of Class I made up 33 3 per cent of the total criticisms offered by the farm boys and only 15 7 per cent of the total criticisms offered by the city boys An additional 20 5 per cent of the farm boys criticisms of mother fell in the miscellaneous class, all of which were called "trivial" in the inventory, while the corresponding percentage of the city boys was only 12.4 Three other classes of criticisms, viz , those concerned with social adjustments, with ideas, attitudes and beliefs, and with work and self-sacrifice, when grouped together accounted for an additional 18 8 per cent for the farm boys and 18 0 per cent for the city boys The remaining 54 per cent of the city boys' criticisms of mother, as compared with only 27 3 per cent for the farm boys, were directed toward (*a*) certain temperamental traits and behavior of the mother which were directly involved in her relationships with the members of her family such as "complains," "finds fault," "inconsiderate," etc., (*b*) her personal habits and conduct such as "talks too much," "smokes," "gossips," etc , and (*c*) her lack of emotional adjustment and control such as "worries," "loses temper," and "nervous" The percentages of farm and city boys' criticism falling into these respective classes were 8 5 com-

TABLE 9

Comparison of the Farm Boys and the City Boys as to Relative Frequency of Criticism of Mothers for Each Class of Criticism

Criticisms of mother, farm boys 117, city boys 89

Group	Per cent of total criticisms of mother	Diff	$\frac{Diff}{\sigma_{diff}}$
	I Matters related to discipline and control		
Farm boys	33.3	17.6	3.03
City boys	15.7		
	II Temperamental traits and behavior		
Farm boys	8.5	—5.0	
City boys	13.5		
	III Personal habits and conduct		
Farm boys	6.0	—17.6	—3.52
City boys	23.6		
	IV Emotional adjustment and control		
Farm boys	12.8	—4.1	
City boys	16.9		
	V Ideas, attitudes, beliefs		
Farm boys	3.4	3.4	
City boys	0.0		
	VI Work, self-sacrifice		
Farm boys	12.0	—0.4	
City boys	12.4		
	VII Social adjustments and behavior		
Farm boys	3.4	—2.2	
City boys	5.6		
	VIII Miscellaneous, trivial		
Farm boys	20.5	8.1	
City boys	12.4		

pared with 13.5, 6.0 compared with 23.6 and 12.8 compared with 16.9

It is apparent, from these figures that the farm boys and the city boys tended definitely to differ as to the sort of behavior, or characteristic which they mentioned as disliked in their mothers Farm boys almost twice as often as city boys mentioned purely trivial matters or more less trivial objections to their mothers' methods of discipline and control City boys, on the other hand, fully twice as often as farm boys mentioned matters pertaining to the personal habits and conduct of their mothers, to their temperamental traits and behavior which affected their relationships with the members of the family and others, or to their emotional adjustment and control.

Undoubtedly some of the items of habit or conduct which were objected to in city mothers, such, for example, as smoking, were relatively uncommon among farm mothers. Perhaps too, because of the relative isolation and fewer social contacts of farm life, the farm boys had fewer, and perhaps less exacting standards by which to judge the personal habits and conduct of their mothers, and hence were less disturbed by them. Consequently, when they were asked to mention something their mothers did which they did not like, the thing that was most likely to occur to them was something, often trivial, connected with their personal relationships with their mothers and which had no particular significance to them in relation to their own personality adjustments. Another possibility is that a cultural difference in general attitude toward parents existed between the two groups. It may be, in more specific terms, that the farm youngster did not feel as free as the city youngster to criticize openly anything serious connected with the conduct or disposition of the parent. Thus, instead of mentioning things about the mother which actually disturbed him to the extent of affecting his own personal adjustments, he tended to substitute trivial matters which were insignificant to him personally.

Summary and Conclusions

The responses of 694 farm, 639 small town, and 545 city adolescents to questionnaire items which asked what their parents did that they did not like, together with their scores on a battery of personality scales were the data of this study. The findings were as follows:

1. Only slightly more than one-third of all the subjects reported things not liked in their parents. The percentages were 35.9 who criticized mother and 26.2 who criticized father. The girls as a group criticized both parents reliably more frequently than did the boys. Of the three residence groups the highest percentage of criticisms come from the small town subjects and the lowest come from the city subjects.

2. Twenty-one significant correlations between criticizing mother or father and the personality scores were obtained. All of these were negative in sign. Fourteen of the 21 came from the city group and 10 of these 14 came from the city boys. Fourteen of the 21

correlations had to do with criticisms of mother. Exactly half of these 14 were contributed by the city boys and not a single one by the farm boys. These seven correlations between criticizing mother and the city boys personality scores ranged up to — .35 for personal adjustment score and average personality rating.

3. When the specific items of parental behavior objected to by each group of subjects were classified according to type or quality, the following group differences were noted: (*a*) Fathers were criticized almost three times as frequently as were the mothers because of personal habits and matters of personal conduct, while mothers were reliably more often criticized for matters related to discipline and control and to their tendency to over-work or to sacrifice self. (*b*) The girls tended more frequently to object to temperamental traits, habits, and difficulties which affected the personal relationships between the parent and others, while the boys tended more frequently to criticize their parents and particularly their mothers because of personal habits and conduct. (*c*) Of the three residence groups, the farm group offered the highest percentage of complaints regarding matters of discipline and control, while the city group gave a reliably higher percentage of criticisms of personal habits and conduct and of temperamental traits and difficulties of their parents. (*d*) When the complaints of farm boys and city boys (the groups that showed greatest contrast as to the number and size of reliable correlations between criticizing and the personality variables) were compared it was found that farm boys almost twice as often as city boys mentioned purely trivial things, or more or less trivial matters pertaining to the mother's methods of discipline, while the city boys more than twice as often as farm boys mentioned personal habits and conduct, temperamental traits and difficulties or matters of emotional adjustment and control. These differences may have been due in part to actual differences in characteristic behavior of mothers in the two cultural situations. They may also be related in part to cultural differences in attitude of boys toward their mothers, that is, either that the farm boys tended to adopt less rigid standards in terms of which they judged their mothers conduct, or that their culturally determined attitudes, or ways of regarding parents tended to inhibit, under the conditions of the experiment, the mention of serious faults or weaknesses in their mothers.

In terms of all the various group comparisons considered in combination, a definite relationship between the "seriousness" of the parental fault objected to and the degree of correlation with personality scores in the subjects was shown. In general, the personal habits, and items of personal conduct, such as "smokes," "drinks," or "swears" appeared to present most difficulty to the youngster. At least for the adolescent boy in the city home situation, parental conduct, and particularly the conduct of the mother was shown definitely to be one of the important factors related to personality development.

References

1. BOWERS, S. M. A study of the child-parent relationship. Ph.D. Thesis, Ohio State University, 1931.
2. GUILFORD, J. P. Psychometric Methods. New York: McGraw-Hill, 1936. Pp. 566.
3. HARTSHORNE, H., MAY, M., & SHUTTLEWORTH, E. K. Studies in the Organization of Character. New York: Macmillan, 1930. Pp. 498.
4. MELTZER, H. Children's attitudes to parents. *Amer. J. Orthopsychiat.*, 1935, 5, 244-265.
5. SIMPSON, M. Parent preferences of young children. *Columbia Univ. Contrib. Educ.*, 1935, No. 652.
6. STEVENS, G. C. Autobiographical material concerning the childhood environments and the effects on after judgments of one-hundred recidivists and one-hundred college freshmen. *Amer. J. Orthopsychiat.*, 1932, 2, 279-303.
7. STOTT, L. H. The relation of certain factors in farm family life to personality development in adolescents. *Nebraska Agr. Exper. Sta. Res. Bull.*, 1938, No. 106.
8. WOODWARD, L. E. Relations of religious training and life patterns to the adult religious life. *Columbia Univ. Contrib. Educ.*, 1932, No. 527.

Nebraska Agricultural Experiment Station
University of Nebraska
Lincoln, Nebraska

The Journal of Genetic Psychology, 1940, 57, 415-428

HOME PUNISHMENT OF ADOLESCENTS*[1]

Agricultural Experiment Station, University of Nebraska

LELAND H. STOTT

That home discipline and punishment are among the factors which condition the attitudes and personal adjustments of children has been shown by a number of studies. Back in 1894, Barnes (1) from a study of the attitudes of 4,000 children toward home punishment decided that "children do not object to any of the ordinary forms of punishment because of the penalty itself," but that they do resent punishment which they feel is unjust. That unjust punishment contributes to juvenile delinquency is suggested by the results of a study by Reinhardt and Fowler (4). They found that approximately twice as many of their delinquent boys as of their non-delinquent boys regarded as the "meanest man" the father who whips his children. Stevens (5) also found that strict home discipline was associated with delinquency and anti-social behavior. That over-strict as well as irregular and inconsistent home discipline is not only related to personality difficulties, but that it is a factor of some importance in relation to subsequent marital unhappiness was shown by Terman in his recent study (9).

In a previous article by the present writer it was shown that among the objectionable items in parents' behavior most frequently reported by adolescent children were those having to do with parental discipline and control (6). The results seemed to indicate, however, that this particular class of criticisms was one of the *least* significant in its relation to the attitudes and adjustments of the subjects. The rural subjects were especially prone to mention matters of discipline when asked what they did not like about their parents, and with them the reliable correlations between mentioning something not liked and their personality test scores were less numerous and smaller than with the other groups. This was particularly noticeable in the

*Received in the Editorial Office on September 18, 1939.

[1]Published with the approval of the Director as Paper No. 247, Journal Series, Nebraska Agricultural Experiment Station. This is one of a series of papers dealing with the results of a research project, the general subject of which was the relation between home-environmental influences and personality adjustments in children.

case of the farm boys' criticisms of their mothers when contrasted with the city boys' criticisms of their mothers

The suggestion was made that perhaps certain of the items of personal behavior which boys find especially objectionable in their mothers, such as smoking, for example, are not so common among farm mothers Or perhaps because of the relative isolation of the farm family where social contacts and relationships are less numerous, the personal habits and traits of the mother do not matter so much to the boy so long as his personal comfort and satisfactions are not interfered with In either case, when asked what he does not like about his mother, items more or less trivial, connected with her methods of control are more likely to occur to the farm boy than to the city boy Another possibility suggested was that there might exist between farm and city youngsters a cultural difference in general attitude toward parents and what they do. Farm children, in other words, might feel less freely disposed to mention items in their parents' behavior which to them are really serious in nature. Whatever may have been the cause, the evidence indicated that from the standpoint of the subjects themselves, criticisms regarding parental control and other personal relationships between parent and child tended to be of little consequence Probably in many cases, such criticisms were not at all correlated with actual punishment and its frequency, or with severity of discipline in the home

The present study, therefore, was concerned with the results obtained from the same group of subjects in response to questionnaire items which asked specifically regarding home punishment These questions were as follows

> Were you scolded or punished at home last week? Yes No
> If so, how were you punished? Whipping
> Slapping
> Scolding
> Other ________
>
> By whom were you punished? .
> For what were you punished? .

The subjects were 1,878 Nebraska high school students ranging in age from 11 to 21 years, of whom 694 lived on farms in the open country, 639 lived in small towns of 600 to 1,300 population, and

545 lived in Omaha. In addition to the home life questionnaire, a number of tests and scales designed to measure various attitudes and aspects of personality adjustment were administered to these subjects. Descriptions of the questionnaire, the scales and the general procedures followed have been reported elsewhere (8).

In administering the questionnaire a special effort was made to obtain the full cooperation of the subjects. They were urged to answer the questions honestly and as accurately as possible, and they were given assurance that the results would be treated completely anonymously.

Results

1. *Frequency of the Report of Punishment*

Of the 1,878 subjects of the three residence groups, 650, or 34.6 per cent, reported that they had been punished at home during the previous week. The boys in slightly but consistently greater proportions, reported punishment. When the percentages of the combined boy groups and the combined girl groups were compared, however, the difference of 3.9 per cent was not significant statistically. These percentages were 36.6 for the boys and 32.7 for the girls.

In Table 1 are given the percentages of the boys and the girls

TABLE 1

Comparison of the Three Residence Groups in Terms of Percentages that Reported Having Been Punished at Home During Previous Week

Group	Per cent of total group reported punished: Farm	Town	City	Groups compared: Farm—Town: Diff	Farm—Town: Diff / σ diff	Farm—City: Diff	Farm—City: Diff / σ diff	Town—City: Diff	Town—City: Diff / σ diff
Boys	29.3	42.9	38.4	—13.6	—3.59	—9.1	—2.39	4.5	1.12
Girls	25.1	38.9	35.6	—13.8	—3.95	—10.5	—2.79	3.3	0.82
Total	27.1	40.7	37.1	—13.6	—5.29	—10.0	—3.75	3.6	1.21

of each residence group who reported in the affirmative concerning punishment. These percentages were greatest for the town subjects and smallest for the farm subjects. The percentage differences of 13.6 and 13.8 between the farm and the town groups and those of 9.1 to 10.5 between the farm and the city groups were quite significant. Those between town and city groups, however, were

not significant. Assuming that the subjects in the three cultural settings were equally conscientious and honest in answering the question, punishment then was less frequent in farm homes than in either town or city homes. This finding is interesting in connection with the fact reported in the previous study (6) referred to above, that the criticisms regarding matters of discipline and control offered by the farm subjects against their parents were relatively great in number in comparison with those offered by the other residence groups.

2. *Relation between Punishment and Personality Adjustments*

As may be seen in Table 2, some very small, but statistically re-

TABLE 2
SIGNIFICANT CORRELATIONS BETWEEN THE REPORT OF HAVING RECENTLY BEEN PUNISHED AND SCORES ON THE PERSONALITY SCALES

	Farm		Town		City	
	Boys	Girls	Boys	Girls	Boys	Girls
Rationality of thinking (Maller "C" score)	—.21				—.27	
Personal adjustment (Maller "A" score)	—.27	—.25	—.29		—.27	—.24
Independence			—.20			
Personal responsibility			—.22		—.21	—.20
Attitude toward work					—.24	
Attitude toward home life	—.23	—.23	—.27		—.21	
Average personality rating[1]	—.23		—.27	—.23	—.30	

[1]The "average personality rating" was the average of the subject's standard scores (modified) on all the personality scales which he completed satisfactorily.

liable correlations were found to exist between recent punishment, as reported by the subjects, and several different personality variables. Correlations greater than —.20 with Maller's personal adjustment were found in all groups except town girls, and with appreciation of home life in all groups except town girls and city girls. Fifteen of the 20 correlations shown in the table came from the boys of the three residence groups. Six of these came from the city boys, five from the town boys, and four from the farm boys. The correlations with "average personality rating" were —.30, —.27, and —.23 for the city, small town, and farm boys respectively. The results of the correlational analysis thus indicate a very slight tendency for the report of recent punishment to bear less relationship

with low personality test scores of the farm subjects than of the city subjects.

3. *Reasons for Punishment*

Seventy-five different reasons for punishment were given by the 650 subjects who reported recent punishment in the home. These reasons are listed in the order of frequency in Table 3. Punishment

TABLE 3

REPORTED REASONS FOR PUNISHMENT AND THEIR FREQUENCIES FOR BOYS AND GIRLS OF EACH RESIDENCE GROUP AND OF THE THREE GROUPS COMBINED, LISTED IN THE ORDER OF THE TOTAL FREQUENCIES

Reasons for punishment	Farm		Town		City		Total	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Getting home late from "date"	11	4	19	22	13	15	43	41
Reason not reported	7	9	13	12	21	12	41	33
Disobedience	9	22	15	11		3	24	36
Nonsensical, trivial things	5	2	3	9	19	9	27	20
Being impudent, "sassy"			3	15	8	8	11	23
Neglecting work	2	1	9	7	9	2	20	10
Getting home late from school	6	3	4	4		4	10	11
Being noisy	1	3	3	5	5	2	9	10
Quarreling	3	2	1	7		4	4	13
Fighting	1	1	3		9	3	13	4
Doing something wrong			7	5	2	2	9	7
Teasing	3	3	2		6	1	11	4
Not being home enough	2	2	1	7	2		5	9
Laziness	4	2	4	1		1	8	4
Neglecting studies	3	1	1	2		3	4	6
Being slow	2	4		2	1	1	3	7
Breaking something	3		3		3	1	9	1
Being cross, "crabbing"		1	1	3		4	1	8
Arguing	2		2	1	2	2	6	3
Low grades	2	2	2	1	1		5	3
Losing temper	1	2		2		2	1	6
Carelessness	3	3	1				4	3
Getting up late	1		2	1	1		4	1
Swearing, foul language		1	2	1		1	2	3
Late to bed		1	1		2	1	3	2
Spending too much money	1		1		3		5	
Not hanging up clothes		3			1	1	1	4
Making mistakes	4	1					4	1
Hitting	2		3				5	
Not following directions					2	3	2	3
Forgetfulness					1	4	1	4
Abusing the family car			5				5	
Rough playing	1		1	1			2	1
Talking back	2	1					2	1

TABLE 3 (*continued*)

Reasons for punishment	Farm Boys	Farm Girls	Town Boys	Town Girls	City Boys	City Girls	Total Boys	Total Girls
Going to too many shows	3						3	
Table manners			3				3	
Not getting enough sleep			3				3	
Staying away too long	1	1					1	1
Misunderstanding	1		1				2	
Disorderly room			1	1			1	1
Neglecting to practice			1	1			1	1
Contradicting					1	1	1	1
Asking questions	2						2	
Playing football	2						2	
Wanting to go some place		2						2
Coming home late from town			2				2	
Eating between meals			2				2	
Doing what I thought was right			2				2	
Trying to discipline me				2				2
Lying				2				2
Attending dances and shows				2				2
Wearing "makeup"				2				2
Doing unnecessary things	1						1	
Going places without permission	1						1	
Not going to Sunday School	1						1	
Taking car to get girl friend	1						1	
Driving too fast	1						1	
Meanness	1						1	
Having a date		1						1
Couldn't tell about party		1						1
Going to football games		1						1
Ignoring a request		1						1
Not giving prompt attention		1						1
Interfering with conversation		1						1
Talking too loud		1						1
Saying something I shouldn't		1						1
"Calling down" younger sister		1						1
"Bawling out" the boys		1						1
Not eating spinach		1						1
Not darning stockings		1						1
Not "putting things in place"		1						1
Breakfast not on time		1						1
Buying a clarinet		1						1
Hunting too late			1				1	
Mischievousness			1				1	
Calling sister names				1				1
Neglecting to see mother				1				1
Total cases of punishment	96	92	129	131	112	90	337	313

was reported but no reason was given by 74, or 11.4 per cent, of the 650 subjects. An additional 47 (7.2 per cent) gave nonsensical or trivial "reasons." None of these are listed individually, but they are grouped together as a single item in the table. The 75 different reasons, therefore, were obtained from 520 individuals.

The most frequently reported single cause of punishment was *getting home late from "date."* Eighty-four (12.9 per cent) of the 650 punished subjects gave this reason. These were nearly equally divided as to sex. The four next most frequent reasons were *disobedience, being impudent or "sassy," neglecting work*, and *getting home late from school*. *Disobedience* and *impudence* were more frequently reported by the girls while *neglecting work* was more frequently reported by the boys. These four reasons accounted for an additional 22.3 per cent of the cases. The remaining 46.2 per cent mentioned the other 70 reasons listed in the table.

When the three residence groups are compared in terms of reported reasons for punishment, some interesting differences appear. Of the 112 city boys and the 90 city girls who admitted recently having been scolded or punished, 18.8 per cent and 13.3 per cent respectively failed to report the cause for punishment. An additional 17.0 per cent of the city boys and 10.0 per cent of the city girls gave very trivial or obviously nonsensical answers to the question. Of the 96 farm boys and the 92 farm girls who admitted recent punishment, only 7.3 per cent and 9.8 per cent respectively failed to report the cause for punishment, and only 5.2 per cent and 2.2 per cent respectively gave nonsensical answers. The corresponding percentages for the town subjects were 10.1 for no answer and 2.3 for noosensical answers from the boys and 9.2 for no answer and 6.9 for nonsensical answer from the girls. Thus, when the two sets of percentages are combined it is found that the city boys gave no actual information regarding reasons for punishment three times as frequently, and the city girls almost twice as frequently as any of the other groups except town girls. For this group the combined percentage (16.1) was reliably smaller than those for the city boys and girls (35.8 and 23.3 respectively).

On the other hand, the remaining 87.5 per cent of the farm boys and 88.0 per cent of the farm girls respectively gave 34 and 38 different reasons for punishment other than those classed as nonsensical, while the remaining 64.2 per cent of the city boys and 76.7

per cent of the city girls respectively gave only 19 and 23 different reasons. The town boys mentioned 35 and town girls 27 specific reasons other than nonsensical.

Being out late at night, as has already been pointed out, was by far the most commonly assigned cause for punishment when all groups were considered. It was the most frequently reported cause for every group except the farm girls. It accounted for 11.5 per cent of the total reports of punishment from the farm boys, 4.4 per cent from the farm girls, 11.6 per cent from the town boys, 11.5 per cent from the town girls, 11.6 per cent from the city boys and 16.7 per cent from the city girls.

Disobedience was the next most common cause for punishment reported by the farm and town groups. With farm boys it came second in frequency, accounting for 9.4 per cent of the total cases. With farm girls it came first, accounting for 23.9 per cent of the cases. With the town groups it came second and third in frequency accounting for 11.6 per cent and 8.4 per cent of the total reports of the boys and girls respectively. In contrast to these groups, in not a single instance was disobedience reported as a cause of punishment by the city boys. It was ninth among the causes reported by the city girls and accounted for only 3.3 per cent of the total cases.

4. *Methods of Punishment*

As for methods of punishment, *scolding* was by far the most frequently reported for both parents by the subjects from all three home settings (Table 4). The percentages of total cases of punishment by parents for each group in which *scolding* was the method used, ranged from 74.0 per cent for the city boys to 86.9 for the town boys. *Made to stay at home* was the next most frequently reported mode of punishment. The relative frequency of this method was greatest for the city boys and least for the town boys. The percentages for the respective groups were farm boys 8.4, farm girls 8.8, town boys 7.4, town girls 11.1, city boys 17.7, and city girls 13.5. *Slapped* was the only remaining mode of punishment reported by all three residence groups. All the percentages were small but the method was reported most frequently by the farm boys (4.3 per cent) and least frequently by the town boys (1.6 per cent). *Whipping* was reported by three individuals of the farm group but not at all by the others.

TABLE 4

METHODS OF PUNISHMENT USED BY PARENTS AS REPORTED BY THE FARM, SMALL TOWN, AND CITY SUBJECTS

Method	Boys: By mother N	%	Boys: By father N	%	Girls: By mother N	%	Girls: By father N	%
Farm subjects								
Scolded	44	46.3	38	40.0	61	67.0	16	17.6
Made to stay home	4	4.2	4	4.2	5	5.5	3	3.3
Slapped	3	3.2	1	1.1	2	2.2	1	1.1
Whipped					1	1.1	1	1.1
Cussed			1	1.1				
"Went out and stayed"							1	1.1
Total	51	53.7	44	46.3	69	75.8	22	24.2
Town subjects								
Scolded	62	50.8	44	36.1	90	70.9	17	13.4
Made to stay home	3	2.5	6	4.9	11	8.7	3	2.4
Slapped	2	1.6			3	2.4		
Privileges taken away	1	0.8	2	1.6				
Made to do work over					1	0.8	1	0.8
Allowance cut			1	0.8				
No supper			1	0.8				
No dates					1	0.8		
Total	68	55.7	54	44.2	106	83.5	21	16.5
City subjects								
Scolded	43	44.8	28	29.2	56	62.9	10	11.2
Made to stay home	7	7.3	10	10.4	9	10.1	3	3.4
Slapped	1	1.0	2	2.1	1	1.1	1	1.1
Allowance stopped	1	1.0	1	1.0			1	1.1
Sent to bed	1	1.0					1	1.1
Talking to							1	1.1
Cutting insults					1	1.1		
(Not reported)			2	2.1	4	4.5	1	1.1
Total	53	55.2	43	44.8	71	79.8	18	20.2

5. *The Punisher*

According to the reports of these subjects, the mothers most often administered the punishment. Of all the cases reported, the mothers administered 64.3 per cent as compared with 31.1 per cent administered by the fathers (Table 5). This relative frequency of punishment by mothers was particularly large with the girls and with the small town group as a whole. Over 80 per cent of the punishment reported by the town girls was administered by the mother as compared with only 16 per cent by the fathers. In the case of the farm

TABLE 5

PERCENTAGES OF TOTAL CASES OF PUNISHMENT WHICH WERE ADMINISTERED BY THE MOTHER, BY THE FATHER, AND FOR WHICH PUNISHER WAS NOT NAMED

Group	Punishment administered by Mother	Father	?
Farm boys	53.1	45.8	1.1
Farm girls	75.0	23.9	1.1
Total farm	63.8	35.1	1.1
Town boys	52.7	41.9	5.4
Town girls	80.9	16.0	3.1
Total town	66.9	28.8	4.2
City boys	47.3	38.4	14.3
City girls	78.9	20.0	1.1
Total city	61.4	30.2	8.4
Total, all groups	64.3	31.1	4.6

boys the percentages were more nearly equal, being 53.1 per cent for the mothers and 45.8 per cent for the fathers.

These results are in fairly close agreement with those reported by Bowers (2). According to the answers of his subjects, 54 per cent of the punishing was done by the mothers as compared with 31 per cent by the fathers. Newell (3), however, reported from his work with juvenile delinquents that more than half of his subjects accused the father of doing most of the punishing at home.

Another interesting variation among the groups of the present study may be seen in Table 5. Sixteen (14.3 per cent) of the city boys who admitted recently having been punished gave no information as to who did the punishing, while only one each of the farm boys, farm girls, and city girls failed to give that information.

Discussion

The data of this study were the reactions of normal adolescent boys and girls to questionnaire items concerning home punishment, which were presented to them in the regular classroom situation. The possible sources of error and inaccuracy in such data are, of course, many. In spite of our efforts to gain their good will and full cooperation, there undoubtedly were among the subjects some who resented being asked such personal questions, and consequently either withheld information or deliberately falsified their responses. Since they were asked to report punishment during the previous

week, errors of memory may have entered. Also, the desire to represent their family life in a favorable light probably varied considerably among the subjects thus predisposing some to "forget" cases of punishment or the reason for punishment and the method used, while others might have tended even to exaggerate in their responses. In regard to methods of punishment, the fact that a few alternative responses were suggested on the blank for the subject to underline, if any of them fit his case, very probably affected the results to some extent. Although "scolding" is undoubtedly the most common mode, yet since it is also one of the mildest forms of punishment, some of the subjects probably accepted the suggestion and underlined "scolding" when otherwise they might have reported some other form. Because of these and possibly other sources of discrepancy between report and actual happening in the home lives of the youngsters, caution must, of course, be exercised in interpreting the results.

Some differences between groups, nevertheless, appear to be worthy of consideration. In the first place, it was found that although the farm subjects more often mentioned matters related to discipline and control as things their parents did which they did not like, reliably smaller proportions of them than of either the town or the city groups reported having recently been punished. These differences are scarcely explainable in terms of differences in conditions, or in procedure of administration of the questionnaire, since the testing of all groups was done by the writer and as nearly as possible under the same objective conditions. The farm and town subjects, furthermore, were members of the same high school groups, the test papers later having been segregated into farm and town "groups" on the basis of residence and father's occupation. The obtained differences must, therefore, represent actual cultural differences apart from the objective test situation.

It is quite possible that scolding and punishment are actually somewhat less frequent on the farm than in town or city. Because of the relative simplicity of the rural family environment with perhaps fewer, and less insistent attractions outside the home which tend to disrupt family unity, it may be that fewer problems of control face parents and fewer causes for punishment arise. On the other hand, farm life, where it is important that each member of the family do his share of the work and assume his share of the responsibility in the common family enterprise, involves the necessity

for rather rigid discipline. The parents and especially the father must direct, generally, the work of the home and farm, and "obedience" on the part of the children is perhaps more frequently and strictly demanded. This presumably would entail considerable scolding and punishment.

These general differences between farm life and life in town or city suggest that there might be associated with them certain cultural differences in attitude toward the whole question of parental discipline and control. Farm people perhaps more generally than city populations, tend to regard the authoritarian relationship between parents and children as the only sort of relationship through which the proper degree of parental control can be maintained. In a previous study (7) some evidence for the existence of such a cultural difference in attitude was presented. Farm parents were found, on the average, to be somewhat more inclined to favor strict parental control over adolescent children than were city parents.

Farm boys and girls, then, may be more inclined to accept without rebellion parental authority, and to take for granted that they should obey their parents because to do so is more a part of the generally accepted pattern in that culture. As was previously suggested the farm youngsters apparently tended not to take too seriously matters of parental discipline and control, perhaps because they were more used to it. They tended to mention such things, rather than the undesirable habits and traits which their parents possessed and which to them perhaps were more serious, when asked what their parents did which they did not like.

Punishment of adolescent age children may actually be less frequent on the farm than in town or city, then, because of the greater tendency on their part to accept the authority of their parents, which tendency has been trained into them from early childhood. The farm subjects of the present study, also because of this general attitude, may have tended more than the others to overlook or forget cases of mild punishment. To be scolded or punished to them was not so frequently, or to the same degree, a humiliating experience. With the town and city subjects, on the other hand, the tendency may have been not to take punishment so submissively or so much as a matter of course. The tendency was, therefore, for them to resent questions about punishment and more frequently to refuse to give information as to who did the punishing or the method used.

Summary

This study was concerned with the prevalence and nature of home punishment of normal adolescent boys and girls of the farm, the small town, and the city home situations, and with the significance of punishment to the youngsters themselves. The results may be summarized as follows:

1. Slightly more than one-third of all the subjects reported having recently been scolded or punished. The report was slightly more frequent from the boys than from the girls. This difference, however, was not reliable.

2. A significantly smaller percentage of the farm subjects than of those from town or city reported recent punishment.

3. A correlational analysis revealed that punishment was associated slightly with low scores on certain tests of personal adjustment and appreciation of home life. In the case of town and city boys punishment showed some relationship with low scores in two varieties of "self-reliance."

4. All together, 75 different reasons for punishment were reported. Over 11 per cent of all subjects who reported punishment failed to state the reason. An additional seven per cent gave obviously nonsensical answers. The great majority of both these groups were among the city subjects. *Getting in late from date* was the most frequently reported reason. It was most common with the city girls and least common with the farm girls. *Disobedience, being impudent or sassy,* and *neglecting work* were the next most frequent. *Disobedience* was much more frequent from the farm subjects than from either of the other groups.

5. Scolding was by far the most common "method" of punishment reported. Other commonly reported ones were *made to stay home* and *slapped*. Three cases of *whipping* were reported, all among the farm subjects.

6. Almost two-thirds of all the punishing was reported as done by the mothers. The percentages of the cases in which the mother did the punishing ranged from 53.7 for the farm boys to 83.5 for the town girls.

The hypothesis that there exists in rural populations a characteristic attitude toward the whole question of parental authority, somewhat different from that which has become common among urban populations, as a result of which rural children more frequently

reach adolescence trained to accept, without conflict or rebellion, the authority and control of their parents, was suggested by the results.

References

1. Barnes, E. Punishment as seen by children. *Ped. Sem.*, 1894, **3**, 235-245.
2. Bowers, S. M. A study of the child-parent relationship. Ph.D. Thesis, Ohio State University, 1931.
3. Newell, H. W. Family attitudes as revealed by the psychiatric examination of 107 juvenile delinquents. *Amer. J. Orthopsychiat.*, 1932, **2**, 377-383.
4. Reinhardt, J. M., & Fowler, V. H. Social and ethical judgments of two groups of boys—delinquents and non-delinquents. *Amer. J. Crim. Law & Criminol.*, 1930, **21**, 364-378.
5. Stevens, G. C. Autobiographical material concerning the childhood environments and the effects on after judgments of one-hundred recidivists and one-hundred college freshmen. *Amer. J. Orthopsychiat.*, 1932, **2**, 279-303.
6. Stott, L. H. Adolescents' dislikes regarding parental behavior, and their significance. *J. Genet. Psychol.*, 1940, **57**, 393-414.
7. ——— Parental attitudes of farm, town, and city parents in relation to certain personality adjustments in their children. *J. Soc. Psychol.*, 1940, **11**, 325-339.
8. ——— The relation of certain factors in farm family life to personality development in adolescents. *Nebraska Agr. Exper. Sta. Res. Bull.*, 1938, No. 106.
9. Terman, L. M. Psychological Factors in Marital Happiness. New York: McGraw-Hill, 1938. Pp. 474.

Nebraska Agricultural Experiment Station
University of Nebraska
Lincoln, Nebraska

The Journal of Genetic Psychology, 1940, **57**, 429-440

STUDIES ON THE REPRODUCTIVE ACTIVITIES OF THE GUINEA PIG. IV. A COMPARISON OF SEX DRIVE IN MALES AND FEMALES[1]

Department of Psychology, Connecticut College

JOHN P. SEWARD AND GEORGENE H. SEWARD

A. INTRODUCTION

The pattern of mating behavior in male and female guinea pigs has already been described in detail (4, 9). In the male, the main features are the active pursuit, mounting, and ejaculation, followed by self-licking of the genitals. Special attention has been paid to the female (7, 8, 21, 22) because of the cyclic morphological changes in the reproductive tract which are related to a behavior cycle. The receptivity pattern is characterized chiefly by the "copulatory reflex," which involves the assumption of a rigid posture with elevation and dilatation of the vulva. This is often either preceded or accompanied by the pursuit and mounting of other animals.

These observations of mating behavior raise the question of its relationship to sexual "drive." In the albino rat several lines of evidence have been brought to bear on this problem by means of the Columbia obstruction box. In normal animals (19) a male drive to females was demonstrated as a function of deprivation interval, while a reciprocal female drive was found to correlate with vaginal cornification. Impotent as compared with potent males (17) have shown a marked inferiority in both crossings and contacts in the obstruction box. In the same direction are the losses in sexual drive following gonadectomy (10). Statistical investigations (2, 3, 16, 18) into the relationship between copulatory behavior and obstruction box performance have yielded correlations ranging from +.32 to +.51. All these results taken collectively indicate a definite relationship, for the white rat, between copulatory behavior and a drive to reach the sex object. The present investigation is

[1]Aided by a grant from the National Research Council Committee for Research on Sex Problems, through the Dept. of Anatomy, Cornell Univ. Med. College. The writers are indebted to Dr. G. N. Papanicolaou of that department for his valuable cooperation.

an attempt to determine whether or not similar relationships hold for the guinea pig

B Method

Behavior in the hurdle box (12), a special adaptation of the obstruction method for this form, was measured under varying conditions of sexual need Male guinea pigs were tested to a female incentive under two conditions: (*a*) immediately after satiation by a receptive female, and (*b*) after several days' deprivation A word is necessary about our use of the term "satiation" The normal male copulates repeatedly until ejaculation occurs, whereupon he abruptly ceases to pay attention to the female for an undetermined period. In our experiments we left the animals together for five minutes of inactivity before removing them By "satiation," therefore, we refer merely to the degree of loss of interest indicated by our criterion

Females were tested in the hurdle box to a male incentive: (*a*) when in oestrus, and (*b*) when not in oestrus By "oestrus" is meant the period of receptivity, which, on the average, occurs every 16 days in this species and lasts approximately eight hours (14, 22) Receptivity was revealed by lordosis, or the copulatory reflex, in response to a male or to stroking with the finger (7, 8, 20) Experimental animals were examined daily for rupture of the vaginal membrane and thereafter were tested frequently for lordosis until a positive response was obtained.

Two series of measurements, *A* and *B*, were obtained with the old and new forms of the hurdle box, respectively. Both boxes, as well as the testing procedure employed with each, have been described elsewhere (12, p 2-6)

1. *Series A*

a Males. Eight normal males were selected on the basis of the apparent strength and persistence of their sexual drive. They varied in age from 121 to 276 days (with one exception, No. 6132, of 531 days) Five of them were tested twice after satiation and twice after deprivation, in *ABBA* or *BAAB* order. The other three were successfully tested once in each condition. All deprivation tests were separated from a preceding satiation by at least six days

The animals used to satiate the males and as incentives were either normal females in heat or spayed females brought into artificial heat by oestrin-progesterone injections (5)

The procedure of a satiation test was as follows The male and female were placed together in a special cage and observed until the male had made no approach to the female for five minutes. The animals were then placed in the hurdle box to test the male's drive to reach the female The deprivation test was preceded by only a few seconds' contact with the female to serve as sexual stimulation.

b. Females The nine females used varied in age from 128 to 140 days. A series of six tests were made in two groups of three each One test (pro-oestrous) was made when the vaginal rupture was first discovered, a second (oestrous), when the female was found to be receptive, a third (postoestrous), when she proved to be no longer in heat The intervals beween these tests varied from 3½ to 38¼ hours, with the average at 14 The other three tests, designed to correspond with the first three, were given in dioestrus, starting from six to twelve days after the preceding oestrus Data are presented from two such series on four animals and one on the remaining five In three cases the pro-oestrous test was precluded by the early onset of heat. In nine of the 13 series the dioestrous runs followed, in four preceded, the oestrous group.

The males used to indicate heat and as incentives were vasectomized, an operation which prevented impregnation without any apparent effect on sex drive [2]

In tests of both males and females the test animal was timed for 10 crossings to the incentive. No crossing was prolonged beyond five minutes, two consecutive uncompleted crossings terminated the test

2 *Series B*

a Males Six males from two litters, 218 and 231 days old, provided a total of 14 pairs of satiation and deprivation tests In this series satiation was determined not only by timing the loss of interest for five minutes but by examining for the vaginal plug

[2]Oslund (11) supplies histological support for this point

To satiate the males we used normal females in heat, as incentives, normal females not receptive but with ruptured vagina

b. Females. Eight females, from 282 to 364 days old, provided data on a total of 27 cycles Only two tests were given in each cycle; one in oestrus, the other in dioestrus between the fourth and the thirteenth day of the cycle Four animals started with a dioestrous test, four with an oestrous. Time of day was controlled by pairing the animals and following each oestrous test on one by a dioestrous test on the other In the intervals between tests they were placed in the hurdle box every few days to maintain familiarity.

Since oestrous in this species is predominantly nocturnal (22) and influenced by light (6), in order to bring our females more conveniently into heat we kept them in a room which was blacked-out by day and illuminated by night Judging by the small percentage of oestrous periods missed, this method proved highly successful.

All hurdle box tests were 10 minutes long except four test pairs on males which lasted 15 minutes There was no time limit on individual crossings

C. Results

1. *Series A*

In this series sex drive was measured by the average time of crossing in a given test Failures to cross within five minutes were arbitrarily scored 300. Tests in which this occurred are indicated by +

a. Males. Table 1 shows the results of 13 tests on males in each condition Nine of the 13 differences are positive, showing faster crossings after deprivation The average difference of 10 56 has a critical ratio of 2 59, which indicates a probability of 03 that it is due to a sampling error.

That the results are not more uniform may be explained by the uncertainty of the criterion of satiation used in this early work Since ejaculation was not verified by examining for vaginal plugs, it is possible that in some cases copulation served to activate rather than satiate the drive (13) This is borne out by two tests on No 6132 not included in the table In the satiation test he crossed in the exceptionally fast average time of 3.1". In the deprivation

TABLE 1

AVERAGE CROSSING TIMES IN SECONDS

Males. *Series 1*

Animal	Test condition: Satiated	Test condition: Deprived	*Diff.* S-D
AII5	26.0	4.9	21.1
No. 6132	18.5	10.0	8.5
EN6	22.2	7.4	14.8
R222	20.1	5.4	14.7
R222	12.2	8.4	3.8
R225	44.4	5.8	38.6
R225	50.4	57.7	−7.3
R129	25.7	2.5	23.2
R129	55.3	37.4	17.9
SA1	22.2	2.3	19.9
SA1	9.8	26.7	−16.9
SA3	6.0	6.4	−0.4
SA3	3.6	4.2	−0.6
Average	24.34	13.78	+10.56
σ Av. D.			4.08
C. R.			2.59

test, after three crossings which averaged 5.7", he deposited a plug in the incentive compartment. The next seven crossings averaged 44".

TABLE 2

ENTRANCES INTO INCENTIVE COMPARTMENT

Males. *Series 1*

Animal	Test condition: Satiated	Test condition: Deprived	*Diff.* D-S
AII5	7	9	2
No. 6132	9	10	1
EN6	5	10	5
R222	3	8	5
R222	10	6	−4
R225	5	7	2
R225	2	8	6
R129	—	—	—
R129	2	9	7
SA1	—	—	—
SA1	0	3	3
SA3	—	—	—
SA3	7	9	2
Average	5.0	7.9	+2.9
σ Av. D.			0.99
C. R.			2.93

Early in the experiment it was noticed that the male, after nosing the door of the incentive compartment, sometimes entered and sometimes remained in the passage-way. Entrances were thereafter recorded and the number of them in each condition appears in Table 2. The data afford additional evidence of stronger drive after deprivation.

b. Females. The crossing times of the females, shown in Table 3, are much slower on the average and more variable than those of the males. The same is true of any phase of the female cycle as compared with either the deprived or satiated condition of the males. This in itself might indicate merely greater activity or agility on the part of the males. When the oestrous times are compared with those in other phases of the cycle, however, no consistent difference appears. The average of all non-oestrous tests is 53.53", almost identical with the oestrous average of 53.64". Inspection of individual scores reveals the same lack of consistency.

TABLE 3
AVERAGE CROSSING TIMES IN SECONDS
Females: *Series A*

	Test condition					
	Pro-		Post-		Dioestrus	
Animal	oestrus	Oestrus	oestrus	1	2	3
R138	——	54.0 +	32.4	7.4	5.4	6.7
R138	——	9.8	66.3	34.8	13.8	——
R149	70.9	127.9 +	125.2	158.7 +	182.4 +	193.6 +
R149	41.2	13.9	86.5 +	60.2	45.9	70.9
R151	123.7 +	144.4 +	300.0 +	15.0	153.1 +	152.3 +
R151	63.3 +	43.5	4.0	59.8 +	60.9 +	76.3 +
R152	52.1	114.3 +	47.8 +	8.7	83.6 +	65.2
R154	——	23.8	17.9	10.3	13.6	49.0
R158	12.5	23.6	168.0	14.8	10.8	24.2
R158	4.9	3.8	3.4	5.3	2.2	2.8
R159	122.0 +	121.2	47.4	89.1	39.1	36.7 +
R160	15.3	13.3	14.2	24.8	13.9	9.3
R170	7.9	3.8	4.0	2.4	2.7	8.6
Average	51.38+	53.64+	70.55+	37.79+	49.80+	57.97+

+ indicates one or more uncompleted crossings scored 300.

The negative character of the results is borne out by the differences in Table 4. In computing the dioestrous-oestrous differences, that dioestrous test was used which corresponded most closely with the oestrous in ordinal position and in time of day. All three of the average differences are approximated by their standard errors.

TABLE 4

AVERAGE DIFFERENCES BETWEEN TEST CONDITIONS

Females *Series A*

Conditions	Av *Diff*	σ Av	C R
Pro-oest—Oest[1]	—9.6	9.4	1.02
Post-oest-Oest	16.9	19.9	0.85
Di-oest-Oest.	—11.5	10.1	1.14

[1] $N = 10$

2. *Series B*

Since the tests in this series had a time limit, the average crossing time was computed by adding the times of all crossings, completed or uncompleted, and dividing by the number of completed crossings. If an animal did not cross at all it was arbitrarily given a score of 600. A second measure of drive, the number of completed crossings, was also used.

a Males. The results on males are presented in Table 5.[8] They confirm the earlier finding (Tables 1, 2) of stronger drive

TABLE 5

CROSSING TIME SCORES IN SECONDS AND NUMBER OF CROSSINGS

Males *Series B*

Animal	Time scores			No. of crossings		
	Satiated	Deprived	*Diff* S-D	Satiated	Deprived	*Diff* D-S
SA5	2.0	3.2	—1.2	—	—	—
SA5	20.4	10.0	10.4	14	21	7
SA5	24.6	23.8	0.8	9	13	4
SA5	24.9	6.5	18.4	11	21	10
SA7	11.3	2.6	8.7	—	—	—
SA7	14.9	8.3	6.6	11	19	8
SA9	17.2	7.5	9.7	—	—	—
SA9	15.4	9.5	35.9	9	20	11
SA15	26.7	8.1	18.6	12	16	4
SA15	142.5	22.3	120.2	4	14	10
SA15	80.2	13.9	66.3	6	15	9
SA17	37.7	6.6	31.1	—	—	—
SA17	8.2	14.1	—5.9	13	18	5
SA19	97.3	14.5	82.8	5	16	11
Average	39.52	10.78	28.74	9.4	17.3	7.9
σ Av			9.77			0.87
C R			2.94			9.08

[8] Number of crossings is shown only in the 10 test pairs of 10 minutes' duration.

after deprivation than after satiation. The difference in number of crossings is more consistent and reliable than in crossing time.

b Females. The results on females are presented in Table 6

TABLE 6

CROSSING TIME SCORES IN SECONDS AND NUMBER OF CROSSINGS

Females *Series B*

Animal	Time scores Diocstrus	Oestrus	*Diff* D-O	No of crossings Dioestrus	Oestrus	*Diff* D-O
SC2	597 0	98 5	498 5	1	4	3
SC2	119 0	600 0 m	—481 0	4	1 m	—3
SC2	118 0	42 6 m	75 4	4	8 m	4
SC4	538 0	99.3	438 7	1	4	3
SC4	600 0	166 3	433 7	0	3	3
SC4	557.0	122.8	434 2	1	4	3
SC4	558 0	560 0	— 2 0	1	1	0
SC6	162.0	99 0 m	63 0	3	5 m	2
SC6	118.8	111 0 m	7 8	4	4 m	0
SC6	125 0	167.7 m	— 42 7	4	3 m	—1
SC6	549 0	271 5	277 5	1	2	1
SC8	54 2	140 7 m	— 86 5	6	3 m	—3
SC8	557 0	31.8 m	525 2	1	8 m	7
SC8	165 7	111 8 m	53.9	3	4 m	1
SC10	168.0	600 0 m	—432 0	3	0 m	—3
SC10	111 8	600 0 m	—488 2	4	0 m	—4
SC10	56 7	600 0	—543 3	7	0	—7
SC10	162 7	600 0 m	—437 3	3	0 m	—3
SC12	128 8	31 9	96 9	4	9	5
SC12	265 0	220 0 m	45 0	2	2 m	0
SC12	538.0	111.6 m	426 4	1	4 m	3
SC12	70 7	45 5 m	25 2	6	8 m	2
SC22	74 4	539 0 m	—464 6	5	1 m	—4
SC22	155 0	295 0 m	—140 0	3	2 m	—1
SC22	46.6	84 0 m	— 37 4	8	5 m	—3
SC22	120.5	260 5 m	—140 0	4	2 m	—2
SC34	44 3	46.9 m	— 2 6	8	7 m	—1
Average	250 41	246 57	3 48	3 41	3 48	0.07
σ Av			62 87			0 62
C R			06			0 11

m=indicates presence of mounting behavior

As in Series *A*, they show a much weaker tendency to cross than do those of the males. Moreover, when oestrous scores are compared with dioestrous, neither times nor crossings show, on the average, any appreciable difference. The 27 individual test pairs are almost equally divided between positive and negative differences

Although two animals, SC10 and 22, have better scores in dioestrous, and two others, SC4 and 12, show an opposite trend, it cannot be said that the animals as a group show self-consistency in this respect.

Especial interest attaches to tests run at a time when the female displayed mounting behavior. Such tests are designated in Table 6 by *m*. When the 19 oestrous tests with mounting are compared with the remaining eight, no difference appears, the number of crossings averaging 3.5 and 3.4 respectively.

D. Discussion

The above results demonstrate that male guinea pigs have a stronger drive to reach a female after several days of deprivation than just after successful copulation. Females display in general a much weaker drive to a male and no stronger when sexually receptive than when non-receptive.

The results on males confirm the similar finding by Warner (19) on the rat in the Columbia obstruction box. The question may be raised, however, as to why the difference we obtained was relatively so small. This may be due partly to the use of a hurdle type of barrier rather than an electric shock. It is also probably due to the less complete satiation afforded our animals. Warner permitted his animals two hours with a receptive female before testing. Stone (15) reports that in male rats the pauses following ejaculation become progressively longer during a three-hour period. Our animals, after one ejaculation, were presumably well above the limit of such a "satiation curve."

Our negative results on females are in contrast to Warner's finding that female rats in the "cornified stage" display a drive equal to the strongest found in males. The contrast is all the more striking in that we used a behavioral criterion of heat. A species difference in sexual behavior is strongly indicated.

With respect to the guinea pig we may adopt either of two interpretations. The more sweeping would be to deny that the female plays any active part in mating. In an early description of oestrous behavior, Loeb (8) cites Miss Lathrop as asserting that females in heat do not go to the male but wait until he comes to them. Such passivity implies either that inner tension does not exist or that it does not impel the animal to seek a sex object.

The phenomenon of mounting behavior, which has been observed

in 90 per cent of oestrous females (21), suggests a more conservative view. We may assume that the female guinea pig possesses a sex drive but that the hurdle box test does not reveal it. We are then called upon to explain why a test which reveals a function in one sex should fail to reveal it in the other.

Were the tests run sufficiently close to the peak of oestrous? Although the intensity of heat reactions was not objectively scaled, the experimenter, after considerable experience in judging varying degrees of heat, selected the animals for testing as close to the peak of excitability as possible. Additional evidence of the intensity of the heat reactions is furnished in the number of females displaying mounting behavior. Young *et al* (21) have supplied statistical data to show that mounting usually coincides with the onset of oestrous, when lordosis is most readily elicited.

Was the sluggishness of the females in the hurdle box due to timidity? Although this possibility cannot be ruled out in all cases, the writers do not consider it a major factor. In Series *B* the animals had been tamed and adapted to the hurdle box for months before the experiment and were selected partly on the basis of their willingness to surmount the barrier. Furthermore, if timidity were a factor, it may be argued that it would tend to make for a superiority of oestrous over dioestrous performance. This is suggested by Anderson's (1) finding that receptive rats were less "emotional" than non-receptive ones.

We are left with the hypothesis that if sexual drive, conceived of as an urge toward a possible mate, occurs in the female guinea pig, its failure to manifest itself in the hurdle box points to some quantitative characteristic of the drive itself. Compared with that of the male, the female drive may be either lower in intensity or dependent on more proximate stimulation. How injections of male hormone would affect the hurdle box performance of female guinea pigs is a further problem of the present series.

E. Summary

Male guinea pigs were tested in the hurdle box to a female incentive after deprivation and after a degree of sexual satiation had been reached.

Female guinea pigs were tested in the hurdle box to a male incentive during the period of oestrous, or sexual receptivity, and during dioestrous.

The males crossed more readily after deprivation than after satiation. The females, however, showed no greater tendency to cross in oestrous than in dioestrous.

References

1. Anderson, E. E. The effect of sexual receptivity upon emotional behavior (timidity) in the female rat. *Psychol. Bull.*, 1938, **35**, 703. (Abstract.)
2. ——— The interrelationship of drives in the male albino rat. I. Intercorrelations of measures of drives. *J. Comp. Psychol.*, 1937, **24**, 73-118.
3. ——— The interrelationship of drives in the male albino rat. II. Intercorrelations between 47 measures of drives and of learning. *Comp. Psychol. Monog.*, 1938, **14**, Pp. 119.
4. Avery, G. T. Notes on reproduction in guinea pigs. *J. Comp. Psychol.*, 1925, **5**, 373-396.
5. Collins, V. J., Boling, J. L., Dempsey, E. W., & Young, W. C. Quantitative studies of experimentally induced sexual receptivity in the spayed female guinea pig. *Endocrinology*, 1938, **23**, 188-196.
6. Dempsey, E. W., Myers, H. I., Young, W. C., & Jennison, D. B. Absence of light and the reproductive cycle in the guinea pig. *Amer. J. Physiol.*, 1934, **109**, 307-311.
7. Ishii, O. Observations on the sexual cycle of the guinea pig. *Biol. Bull.*, 1920, **38**, 237-250.
8. Loeb, L. The correlation between the cyclic changes in the uterus and the ovaries in the guinea pig. *Biol. Bull.*, 1914, **27**, 1-44.
9. Louttit, C. M. Reproductive behavior of the guinea pig. I. The normal mating behavior. *J. Comp. Psychol.*, 1927, **7**, 247-263.
10. Nissen, H. W. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat. *Genet. Psychol. Monog.*, 1929, **5**, 451-547.
11. Oslund, R. M. A study of vasectomy on rats and guinea pigs. *Amer. J. Physiol.*, 1924, **67**, 422-443.
12. Seward, J. P., & Seward, G. H. Studies on the reproductive activities of the guinea pig. I. Factors in maternal behavior. *J. Comp. Psychol.*, 1940, **29**, 1-24.
13. Seward, J. P. Studies on the reproductive activities of the guinea pig. III. The effect of androgenic hormone on sex drive in males and females. *J. Comp. Psychol.*, 1940, **30**, 435-449.
14. Stockard, C. R., & Papanicolaou, G. N. The existence of a typical oestrous cycle in the guinea pig—with a study of its histological and physiological changes. *Amer. J. Anat.*, 1917, **22**, 225-265.
15. Stone, C. P. Temporal relationships in the copulatory acts of adult male rats during three hour tests. *Psychol. Bull.*, 1939, **36**, 535. (Abstract.)

16. Stone, C. P., & Barker, R. G. Spontaneous activity, direct and indirect measures of sexual drive in adult male rats. *Proc. Soc. Biol. & Med.*, 1934, **32**, 195-199.
17. Stone, C. P., Barker, R. G., & Tomilin, M. I. Sexual drive in potent and impotent male rats as measured by the Columbia obstruction apparatus. *J. Genet. Psychol.*, 1935, **47**, 33-48.
18. Stone, C. P., Tomilin, M. I., & Barker, R. G. A comparative study of sexual drive in adult male rats as measured by direct copulatory tests and by the Columbia obstruction apparatus. *J. Comp. Psychol.*, 1935, **19**, 215-241.
19. Warner, L. H. A study of sex behavior in the white rat by means of the obstruction method. *Comp. Psychol. Monog.*, 1927, **4**. Pp. 68.
20. Young, W. C., Dempsey, E. W., Hagquist, C. W., & Boling, J. L. The determination of heat in the guinea pig. *J. Lab. & Clin. Med.*, 1937, **23**, 300-303.
21. ———. Sexual behavior and sexual receptivity in the female guinea pig. *J. Comp. Psychol.*, 1939, **27**, 49-68.
22. Young, W. C., Dempsey, E. W., & Myers, H. I. Cyclic reproductive behavior in the female guinea pig. *J. Comp. Psychol.*, 1935, **19**, 313-335.

Department of Psychology
Connecticut College
New London, Connecticut

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1940, 57, 441-444

A TEST OF EQUIVALENCE OF AN EMOTIONAL STIMULUS IN A MACACUS MONKEY*

Department of Psychology, Brooklyn College

CHARLES NILSON WINSLOW[1]

The subject of this experiment was a macacus female monkey, about one and a half years of age at the beginning of observations, and two and a half years of age at the time when they were completed. She was already sexually mature and well developed physically in other ways, and during the entire period appeared to be in excellent health. The difference between the responses of the animal to men and women became apparent a few days after she had been acquired by the laboratory. She was friendly and cooperative with men, and during the period of confinement in the laboratory did not bite a man nor show any other sign of animosity. The response to women was in striking contrast. An intense emotional response was elicited by the approach of any woman to the cage. The response was labelled "anger" by the writer and by every other person who observed it. The difficulty of differentiating between the emotional responses of anger and fear was recognized by the experimenter, but the aggressive character of the response instead of withdrawal, served as the basis for classification.

The response to women could be described as follows: Immediately upon the approach of a woman the monkey began to jump up and down in the cage, leap to the metal grating of the door, cling to it, shake it, and frequently reach the front paws through the meshes to snatch the woman's clothing, if it was within reaching distance. These locomotor responses were usually accompanied by clearly audible throaty growls. Perceptible variations in the intensity of

*Received in the Editorial Office on August 17, 1939.

[1]The writer wishes to acknowledge his indebtedness to Mr. Jack Kronenberg for his valuable assistance in performing the tests.

the response occurred, even from one woman to another, but the various features of the responses as just described appeared invariably.

After a year in the laboratory, a length of time judged to be sufficient to provide for adaptation or toleration to women, since the monkey might conceivably have developed this response on her first contacts with women, the character and the intensity of her misogynist reactions remained unchanged. The author, thereupon, prepared several tests to be used with the monkey to determine the characteristics of the human female perceptual pattern which elicited the above described emotional response. Kluver (2) has succinctly stated the problem of equivalence in the sensory field and the way in which it can be tested as follows:

> Notwithstanding certain changes in the stimulus-situation, the stimuli are in some way "identical" or "similar"; they call forth the same "response", they are from the point of view of the reaction produced "equivalent." It is possible, of course, to introduce changes which destroy this "equivalence", upon the introduction of such changes the stimuli become "non-equivalent."

The writer in this experiment defined as an "equivalent response" to an "equivalent stimulus" the occurrence of the same pattern of emotional reactions which was always elicited in response to women. The absence of any aspect of this response pattern would invalidate the claim for equivalence.

The first test consisted of presenting a large boudoir doll 28 inches tall, with a long draped dress, and hair and features that were fair substitutes of the human form. The doll elicited a response of curiosity in the monkey. When the doll was placed in front of the cage door, the monkey reached its fore paw through the meshes and touched and manipulated the fabric of the doll's dress. The real aspects of the "equivalent" emotional response were absent, however.

In a second test a man (height 5 feet, 9 inches) was dressed in a woman's costume complete from hat to shoes. Efforts were made to duplicate a woman's body form by providing breasts and putting on a cosmetic "make-up" of lipstick and rouge. To the man in natural form and attire the monkey's response had been one of friendly indifference, but when he appeared in the feminine disguise, she became markedly interested. The click of the high heel

shoes on the cement floor was the center of much of this interest, and she followed the shoes continuously with her eye movements. When he stood in front of the door of the cage, she took the dress fabric in her fore paws, handling it, but occasionally pulling at it angrily, and at the same time jumping up and down. It seemed, therefore, that certain elements of the emotional response pattern had been evoked by the womanly disguise, but it could not be said that "equivalence" had been reached.

A third test consisted of testing the response to a woman (5 feet, 7 inches tall) dressed in a man's clothing. The monkey's response to her in her usual attire was the complete one of anger previously described. One-half hour before the change of clothing occurred, a test for the emotional response to the woman revealed that it was present in every detail. The woman then dressed in a man's clothes from hat to shoes, putting the hair up to show the ears, and removing all cosmetics. The entrance of the disguised woman produced parts of the emotional response in the monkey, but only after about 20 seconds. Then the jumping up and down and the reaching through the meshes of the cage door to snatch at the clothing occurred. The throaty growls of anger were never made, however. Thus, it would seem that the response had been modified by the absence of parts of the "equivalent" stimulus pattern of a woman. The fact that the parts of the response which did occur appeared after a 20-second delay indicates that the woman at first was almost a changed perceptual pattern through her disguise. The non-male part or parts of the pattern, possibly the bodily odor, emerged, and with them a change in the meaning of the stimulus.

The woman in her usual attire appeared before the monkey on the next day, with the result that the emotional response occurred in its usual form.

Throughout the tests with the humans, the voices were eliminated as stimuli, since they could hardly be sexually interchanged. This aspect of the perceptual pattern and that of the sexual odor were the parts of the natural stimulus that were missing, and their presence may have been required if an "equivalent" stimulus was to be produced, as judged on the basis of an "equivalent" response.

References

1. Kluver, H. The equivalence of stimuli in the behavior of monkeys. *J Genet Psychol*, 1931, 39, 3-27
2. ——— Behavior Mechanisms in Monkeys. Chicago: Univ. Chicago Press, 1933. Pp 387

Department of Psychology
Brooklyn College
Bedford Avenue and Avenue H
Brooklyn, New York

The Journal of Genetic Psychology, 1940, 57, 445-450.

THE CONSTANCY OF PSYCHO-MOTOR TEMPO IN INDIVIDUAL INFANTS*

Clinic of Child Development, Yale University

LOUISE BATES AMES

A. INTRODUCTION

A quantitative study of motor speeds in varied behavior situations has suggested considerable consistency in the tempo of movement of any one child. Children who creep slowly, climb and prehend slowly. The characteristic tempo of movement does not appear to change markedly as a result of practice.

Data for the present study consisted of approximately 4,800 feet of 16 mm cinema records of the manipulatory and locomotor behavior of eight infants, recorded over a three-year period. Approximately 400 feet of film picturing the manipulatory behavior involved in combining objects were available for each subject, or about 56 feet at each monthly age level. About 100 feet each of creeping and climbing behavior were available for each subject. Since the camera was set at a standard speed of 16 frames per second (or 40 frames per foot) duration of movement could be measured by counting frames.

Tempo of movement was determined for individual subjects at each age level. The data were then analyzed to throw light on the following questions: (*a*) Did each child exhibit a constant tempo for different kinds of behavior, that is, was he consistently the fastest, slowest, etc., or did he vary in speed from one behavior to another? (*b*) Was tempo constant from age to age, or did it speed up appreciably with age and experience?

B. CONSTANCY OF TEMPO IN DIFFERENT KINDS OF BEHAVIOR

When the subjects were arranged in rank order with regard to speed of performance in several different kinds of behavior, the psycho-motor rate for any one individual proved to be so typical and constant that the tendency was for the rank order to remain the

*Accepted for publication by Arnold Gesell of the Editorial Board, and received in the Editorial Office on September 13, 1939.

same for each of the kinds of behavior (4, p 222) Thus, as a rule, the child who crept fastest, prehended fastest He also attained each stage in the creeping sequence soonest. Not only were his specific behaviors rapid, but also his rate of development. The child who crept most slowly prehended most slowly, and attained each of several stages of prone behavior last A similar constancy of tempo was observed in the other subjects

The following table lists eight infants in rank order with regard to *creeping speed, age of attaining the sequential stages of prone behavior, speed of combining pellet and bottle,* and *speed of combining cup and spoon* The usual order with regard to speed of behavior is *G*-44[1], *G*-48, *B*-38, *G*-47 (when data are available), *G*-31, *B*-36 *G*-43 varies in position, and *G*-31 combines cup and spoon the most quickly, though she maintains next to the last position in the other behaviors This indicates considerable constancy in tempo, since the mathematical chances of the same order occurring for five cases in three behaviors are only 1 in 14,400; the chances for four cases maintaining the same order for four behaviors are 1 in 13,824.

C CONSTANCY OF PSYCHO-MOTOR TEMPO FROM AGE TO AGE

The speed of single limb movements and of patterned combinations

TABLE 1

RANK ORDER FOR CASES ARRANGED ACCORDING TO SPEED OF BEHAVIOR[2]

Creeping speed[2]	Prone, age of attaining stages	Pellet	Cup and spoon
(*G*-43)	(*G*-43)		(*G*-31)
G-44	*G*-44	*G*-44	*G*-44
(*B*-34)	(*B*-34)		
G-48	*G*-48	*G*-48	*G*-48
B-38	*B*-38	*B*-38	*B*-38
		(*G*-47)	(*G*-47)
		(*G*-43)	(*G*-43)
G-31	*G*-31	*G*-31	
B-36	*B*-36	*B*-36	*B*-36

[1]Subjects are designated as *G*-43, *B*-38, etc, as in Gesell, A, *et al*, *An Atlas of Infant Behavior*, New Haven Yale University Press, 1934, Vol I, where these same subjects are pictured in these and other behavior situations

[2]Data for all subjects are not available for every situation Creeping speed is based on timing for movement of each of the four limbs in that stage of the normal prone sequence when the child creeps on hands and knees

of limb movements was ascertained in *creeping* behavior. In the course of development a pattern of progression may elaborate or be partially replaced by a more complicated pattern, but if a certain movement is retained as a component of the pattern it remains almost the same in form, and the durational value both of a component and of a total pattern remains highly constant. This constancy is illustrated in Table 2, which gives time values for the "near step" stage (1, p. 433) of the creeping behavior of Boy *D*.[3]

TABLE 2
TEMPORAL DESCRIPTION OF CREEPING BEHAVIOR, BOY *D*

Limb moving	Duration of movement in seconds 40 weeks	44 weeks	48 weeks
Left hand	.34	.34	.36
Right knee	.56	.67	.59
Right hand	.34	.34	.34
Left knee	.25	.28	.31
Time between two sets of limbs	.03	.06	.03
Total time for pattern[4]	1.00	1.06	.90

There are marked individual differences in the total time for a given creeping pattern at any one age. For example, four infants at 48 weeks of age perform the total four-limb pattern for the "near step" stage in .90 seconds, 1.06 seconds, 1.23 seconds, and 2.22 seconds respectively. But the speed of movements, whether of one limb or of the total pattern is significantly constant for any given child.

Stair climbing behavior (2) was similarly measured in 12 subjects who were observed from 52 to 80 weeks and in some cases to two and three years of age. Cinemanalysis showed that from the time when climbing is first definitely exhibited, it retains a characteristic form with but slight increase in speed over a period of months, until the child eventually walks upstairs. The following figures (Table 3) describe temporally the four-limb stair-climbing behavior of a normal infant from the time when it is first till the time when it is last exhibited.

Time determinations of *manual behavior involved in the combination of objects* were made in three of the normative situations com-

[3]Locomotor behavior of Boy *D* is pictured in Gesell, A., *et al.*, *An Atlas of Infant Behavior*, Vol. II. Records were taken on 35 mm. film.

[4]Total time for pattern is less than the sum of the times for separate limbs since two of the limbs are usually moving simultaneously.

TABLE 3
TEMPORAL DESCRIPTION OF STAIR-CLIMBING

Limb moving	Duration of movement in seconds 56 weeks	80 weeks	2 years	3 years
Left foot	93	75	62	62
Left hand	56	56	56	56
Right foot	1 25	75	68	68
Right hand	56	56	50	50
Left foot	93	75	62	62
Whole pattern	2 62	2 31	2 00	1 81

monly used in the Yale developmental examinations, namely: cup and spoon, cup and cube, and pellet beside bottle. These situations were selected for two reasons: first, because the conditions of the normative examination guarantee a standard presentation of stimuli from child to child and from age to age; second because though the component reactions are simple and direct, they involve the more complex pattern of combining objects. Combining time was reckoned from the moment of lifting an object (pellet or spoon or cube) from the table, to the moment when this object was held directly over the object with which it was combined (bottle or cup). Time of release was not measured, as this is a more complex act, not occurring at all in the first combinations.

Data for cup and spoon, and pellet beside bottle are given in tabular form. Data for cup and cubes being less complete are merely listed. The figures indicate that for these three situations there is very little decrease in speed (that is, quickening of movement) with increased age and experience.

TABLE 4
CUP AND SPOON BEHAVIOR

	Time in seconds for combining spoon and cup 44 weeks	48 weeks	52 weeks	56 weeks	60 weeks
G-31	76	76	82	.76	—
G-44	69	82	.88	—	—
G-48	—	—	88	—	82
B-38	—	1 07	1 20	1 20	—
G-47	—	1.07	1.20	1 26	1 00
G-43	95	1 20	1 13	1 32	95
B-36	—	—	1 07	1 39	1 00

TABLE 5
PELLET BESIDE BOTTLE

	Time in seconds for combining pellet and bottle							
	48 weeks	52 weeks	56 weeks	60 weeks	80 weeks	2 years	3 years	4 years
G-44	—	.50	—	—	.76, 50*	76	69, 57*	.63*
G-48	—	—	76	—	76	76	63	69, 63*
B-38	69	—	—	—	69	82	69	—
G-47	—	88*	88	95	—	95	76, 69*	—
G-43	—	—	95	76	1 00	95	88	76
G-31	—	1 07	1 20	—	1 00	—	—	—
B-36	—	—	—	1.39*	1 26	95**	1 07 **	95

*Left hand
**Same time with each hand
Otherwise, right hand.

TABLE 6
CUP AND CUBES

G-43 (48 to 60 weeks)	Increases from 44 seconds to	50 seconds
B-38 (48 to 80 weeks)	Decreases from 88 seconds to	69 seconds
B-34 (44 to 56 weeks)	Decreases from 95 seconds to	82 seconds
G-47 (44 to 60 weeks)	Decreases from 1 07 seconds to	88 seconds
B-36 (52 to 80 weeks)	Decreases from 1 13 seconds to	1 07 seconds

D. SUMMARY AND CONCLUSION

1 Cinema records of the locomotor and manual behavior of eight infants were analyzed to determine the constancy of speed of movement

2 Cinemanalysis revealed that in creeping and stair-climbing, time for any single limb movement as well as time for the total four-limb pattern did not appreciably decrease with age and prolonged use of the pattern Time for combining one object with another also remained constant in any one child from age to age

3 When the eight subjects were arranged in rank order with regard to speed of combining pellet with bottle and cup with spoon, creeping speed, and time of attaining successive stages of prone behavior, four retained the same relative rank order for each type of behavior Two more maintained a constant rank order for those situations on which data were available Only two subjects varied in position

The fact that in such various fields of behavior new and often complex behavior patterns can appear for the first time in a form

so complete that several weeks, often several months, of exercise do not appreciably change either their form or speed, strongly suggests that such patterns are determined by internal maturational rather than experiential factors. Children do simpler things first and more complex ones later, but the individual movements, whether of simple or advanced patterns, remain remarkably constant.

References

1. AMES, L. B. The sequential patterning of prone progression in the human infant. *Genet. Psychol. Monog.*, 1937, 19, 409-460.
2. ——— Some relationships between stair-climbing and prone progression. *J. Genet. Psychol.*, 1939, 54, 313-325.
3. GESELL, A., *et al.* An Atlas of Infant Behavior: A systematic delineation of the forms and early growth of human behavior patterns, illustrated by 3,200 action photographs. New Haven: Yale Univ. Press, 1934. Vols. I, II.
4. GESELL, A. (assisted by AMES, L. B.) Early evidence of individuality in the human infant. *Sci. Mo.*, 1937, 45, 217-225.
5. GESELL, A., THOMPSON, H., & AMATRUDA, C. S. The Psychology of Early Growth. New York: Macmillan, 1938. Pp. ix+290.
6. GESELL, A. Prone Progression in the Human Infant. From the Photographic Library of the Yale Clinic of Child Development, 35 mm. film, 1,000 feet; 16 mm., 400 feet.

Clinic of Child Development
Yale University
14 Davenport Avenue
New Haven, Connecticut

APPARATUS

The Journal of Genetic Psychology, 1940, 57, 451-460

THE CONSTRUCTION AND USE OF A NEW APPARATUS IN THE TRAINING OF THE RAT IN AUDITORY DISCRIMINATION PROBLEMS*

Department of Psychology, University of Illinois

W. A. HUNTER AND L. A. PENNINGTON

Several investigators have reported from time to time considerable difficulty in training lower mammalian forms, particularly the rat, to respond differentially to auditory stimuli. Still greater difficulty has been encountered in the attempts made to train these animals to localize the direction of sounds and to discriminate pitches and loudnesses. Experimentalists are in general agreement that in the investigation of problems related either to the peripheral or to the central aspects of the auditory modality one of the major problems is the development of a methodology for training the infra-primate subject. Four difficulties relative to the available procedures in this field of investigation can be found in the literature. First (6), certain methods require prolonged training periods and so preclude the formulation of any very detailed operative program. Second (1), other methods of training eventuate in such decidedly abbreviated responses to auditory stimulation that it is difficult, if not upon occasion impossible in the later stages of learning, to determine whether the animal has or has not responded. Third (7), some experimental devices are so complicated that they require considerable time for the adjustment of the animal in the apparatus as well as a more or less extended preliminary training procedure. Fourth (3), the control of the motivation factor in animal experimentation is an ever present problem in the training procedures. In problems using food as an incentive, by way of illustration, it is

*Received in the Editorial Office on August 17, 1939.

not known whether motivation is constant from one experimental animal to another. Further, the use of this incentive entails considerable time and difficulty in its regulation. In contrast, however, evidence (5) indicates that electric shock as an incentive offers a source of motivation which may require a minimum of time and difficulty in its regulation and control. The incentive, for instance, can be maintained at a constant level of intensity throughout a daily training series. It can also, if desired, be increased within limits as the experiment progresses and can, therefore, be relied upon to maintain a uniform level of performance in the experimental animal. Finally, and as a corollary of the fourth factor just mentioned, the "reality" of the problem for the animal may be raised. Although no experimental animal can be expected to introspect for the experimenter, it is conceivable that the behavior of an animal in a shock-escape arrangement may well show fewer "random" reactions than does the animal in a maze situation. Thus the experimenter may infer from the observations of his animal that the problem for the subject in the shock-escape arrangement corresponds more closely to the problem formulated for the subject by the experimenter than does the reaction of the animal in some modification of a maze discrimination arrangement (6). That this factor, if our assumption is correct, might well have some differential bearing upon the inferred effect of an independent variable on the behavior which is being measured should not be ignored. It follows, therefore, that it is desirable to have the animal reacting as consistently as possible to the specific problem established by the experimenter. This is particularly true if the experimenter is to use these reactions (such as error scores) as measures of the rôle of the independent variable in the acquisition of a given response. The nature of the training apparatus for use in sound experiments, therefore, should present to the animal a situation within its repertoire of activities and one in which its responses can legitimately be inferred to be measures of habit acquisition. For these reasons, then, it has been found advisable to construct an experimental arrangement for use in comparative studies in audition to fulfill the requirements just enumerated. The purpose of this paper is to describe in detail this apparatus and to indicate the nature of the findings obtained from its use during a series of preliminary experiments.

Description of the Apparatus

The details of the construction of the experimental arrangements, described only briefly in a previous article (2), are presented below. The design and construction of this box, it may be said at this point, can easily be varied to meet the requirements of numerous comparative studies in audition. Only the essential features of this apparatus, however, will be described at this time.

The apparatus may be described as follows. It is composed of four units which simplified the moving and cleaning of the arrangements. The first unit consists of a floor grid two feet square. This grid has been constructed of 3/16 inch round steel rods mounted in rubber tubing at the ends and center. These rods are spaced at ½ inch intervals and are stapled to a substantial wooden frame. This arrangement of the rods insures fairly consistent self-cleaning in as much as droppings fall through the spaces and thus prevent, for the most part, shorting in the grid. Second, the walls of the cage, constructed in one unit, are the same length and width as the grid. These walls rest upon, but are insulated from, the grid by blocks of sponge rubber at each corner of the apparatus. The walls are constructed of galvanized iron and are securely fixed to a wooden frame-work. The use of this wall material has prevented any wall clinging shock-escape responses on the part of the experimental animal. The walls are 16 inches in height and generally prevent, therefore, the animals from attempting to jump and cling to the cage cover. Third, a piece of ¼ inch wire mesh on a wood frame serves as a cover for this modified cage. A piece of plate glass, however, has been found to be more satisfactory as a cage cover in those instances in which the animals attempt to jump and to cling to the wire platform during the preliminary training procedures. Fourth, a small platform is located in the center of the floor grid. This platform serves as a second electric grid. It has been constructed of wood and measures 2 x 5 ins. in length and width. Its height is three inches. Across each end of the block holes were drilled one-half inch apart. Into each of these openings was inserted a glass tube two and one-half inches long. Strips of copper plate about seven inches long with flanges so cut that each would fit into a glass tube inserted at each end of the block formed the grid. The shock incentive is supplied by passing the 110-volt alternating current from the light source through 360,000 ohms resistance. This current of 3 milli-

amperes supplied to the grid has been found to be of sufficient intensity to cause all animals used in this series of preliminary experiments to move about the cage The animal can thus be trained to acquire an adaptive shock-escape response when this current is supplied to the grids, even though the variation in skin resistance from animal to animal undoubtedly makes the effective incentive shock smaller for some animals than for others

Training Procedures

Following the construction of the apparatus two preliminary studies were carried out in order to test the efficiency of the experimental arrangements The first study used a 1,000-cycle tone and the second a sound from a buzzer as acoustic stimuli. The training procedure in both studies consisted of three stages. The first stage consisted of a period during which the animals were given an opportunity to explore the training apparatus During the second stage the experimental animal was trained to escape the shock stimulus by moving from one grid to the other. In the third stage the rat was trained to avoid the electric shock by making the just acquired shock-escape response to an acoustic stimulus

The three steps in the training of the experimental animals are described below in greater detail First, the animals selected for this series of experiments were placed, by small groups, in the apparatus for a period of at least two hours during the first day of the training procedure It was hoped that this period would serve to accustom the rat to the apparatus, and consequently, to decrease the amount of exploratory behavior which might well occur between the successive presentations of the stimulus On the next day the animals were placed, one at a time, in the experimental apparatus At this time each was trained to escape the shock stimulus by jumping upon or down from the small platform The response necessary to escape the shock stimulus of course depended upon the position of the rat in the apparatus at the time the current was applied In order to elicit the jumping response, the electric current was presented continuously to the grid on which the rat happened to be until the animal moved to the other grid As a precaution against establishing periodic responses to temporal intervals, the time between the completion of one response and the stimulus for the subsequent response was varied at random between 20 to 40 seconds The

number of presentations of the shock stimulus varied with the facility with which the escape response was acquired. The criterion for the acquisition of this preliminary shock-escape response was arbitrarily set at four successive adaptive responses each occurring within 10 seconds after the onset of the shock stimulus. Third, on the day following the completion of the preliminary training the actual experimental program was initiated. In this stage it was necessary for the rat to learn to avoid the shock. The individual animal was placed in the apparatus where it received the uninterrupted auditory stimulus until the shock avoidance response was made within the arbitrarily chosen 10-second interval. If, however, the rat made the adaptive response any time within this 10-second interval the auditory stimulus was discontinued. If, on the other hand, the adaptive response was not elicited by the end of this 10-second period, the sound stimulus was discontinued and the shock presented continuously until the jumping response was made. The time between one adaptive response and the presentation of the auditory stimulus for the next trial was so varied from 20 to 40 seconds that the formation of a trace response was precluded. The number of presentations of the stimulus per day of the training series varied according to the speed and consistency with which the animal reacted to the tone or to the buzz. The number of presentations of the stimulus varied from 9 to 20 per day. In case the animal made 9 successive adaptive responses at the beginning of the daily series, or made 9 adaptive responses out of 10 successive presentations, the daily series was considered completed. The criterion of learning was thus considered to have been reached for that day's training series. The training was resumed on the following day for all experimental animals until each had reached the arbitrarily selected criterion of 90 per cent accuracy over a period of three successive days. In order to eliminate increasing delays (previously observed in the preliminary training series) in shock avoidance responses to the sound stimulus shock reinforcements were given. Two of these reinforcements were always presented at the beginning of the daily training series. Others were presented whenever the animal was observed to be responding consistently toward the close of the 10-second interval of auditory stimulation.

Experimental Tests

In the present section the detailed description of the two studies

mentioned above will be discussed. The purpose of each has been from the point of view of this paper to test the efficiency of the apparatus already described.

1. *Experiment I*

A group of 31 albino rats each about 90 days of age were trained to escape electric shock by jumping down from or upon the small grid platform in response to a 1,000-cycle tone generated by an audio oscillator. In Table 1 are found the data collected during

TABLE 1

TRAINING DATA OBTAINED FROM NORMAL AND OPERATED ANIMALS

	No.	Days	Trials	*SD*	Errors	*SD*	Reliability
Normal	17	17	325.7	95.9	171.5	78.7	.95±.16
Operate	14	20	402.0	179.0	248.0	157.2	.97±.17

this investigation. It is to be noted that 17 of the 31 animals were trained prior to cortical operation, while the remaining 14 were trained initially following cerebral ablation. It should be noted, likewise, that the trial and error scores presented in Table 1 are averages. Although too few cases have been studied to justify the statistical treatment used, the results do indicate certain trends. For this reason these data have been included. Reference to the coefficients of reliability, as an illustration, indicates a trend toward a high degree of consistency. These coefficients have been calculated on the basis of the error scores made during the first half and the last half of the training series for 14 animals in each group. These coefficients were then corrected by the Spearman-Brown "prophecy formula" with the results which have been included in the table. It is believed, therefore, that the data tabulated on the basis of the 17 normal and 14 operated animals indicate that this method described in the preceding paragraphs is rapid and reliable for both groups of subjects studied.

2. *Experiment II*

A group of 42 normal albino male rats between two and three months of age and a group of 13 operated male albino rats four months of age were trained to avoid an electric shock by jumping onto or down from the small grid platform of the apparatus in response to buzzer stimulation. In Table 2 are found data which

TABLE 2

Training Data Obtained from Normal and Operated Animals

	No.	Days	Trials	*SD*	Errors	*SD*	Reliability
			Initial learning scores				
Normal	42	4.6	70.6	29.3	26.6	19.2	.90±.03
Normal	13	5.7	94.5	46.7	42.6	29.4	.95±.03
			Retention test scores				
Normal	10	3.0	30.8	2.0	1.9	1.1	

indicate the speed of learning as well as the degree of retention of the auditory habit by normal animals. These data are paralleled by data obtained from animals with cortical lesions. The reliability of the apparatus, obtained by correlating the errors for odd and even trials, is indicated for both the normal and operated animals. The data indicate, it is maintained, that when a buzzer is used as a stimulus generator the auditory habit is acquired in a relatively short time by both normal and operated animals. Further, the data show that the habit is retained by normal animals during a 10-day rest period with little or no loss.

Upon the acquisition of this simple adaptive response to the 1,000-cycle tone or to the buzzer a series of control experiments were carried through to check the existence of possible secondary cues. Changes in the location of the stimulus generator both as to distance and direction within certain obvious limits failed to disturb the accuracy with which the trained animal responded. Alterations in the illumination within the apparatus as well as external to it failed to alter the accuracy with which all trained animals responded to the auditory signals. Vibratory cues were checked by cutting the vibrissae and by removing the sound generator from its usual position. No behavior changes were observed. Possible secondary cues, such as clicks from the shock giving devices, were all tested with negative results. The behavior of the animal was studied, in addition, in the absence of the auditory stimulus to determine whether the adaptive response would occur in an anticipatory manner. In the absence of auditory stimulation the animals, trained to respond to a specific auditory signal, failed to respond in a sequential anticipatory temporal order. It has been concluded, therefore, that the present apparatus and the training procedures used in conjunction with it have controlled the rôles of the secondary cues mentioned above.

In addition to the control experiments already described one other

was carried out It is conceivable, it may be argued, that the animal might after preliminary training adapt to the experimental situation by numerous jumps upon or down from the small grid These reactions, occasionally coinciding with the presentation of the auditory stimulus might then be in response to other stimuli, or to chance factors, instead of to the specific auditory stimulus In order to check upon this possibility the following test was made. Late in the training period, and, thereby giving the advantage to the hypothesis about to be tested, each animal of the two groups was given two tests One of the tests preceded and the other followed the presentation of the 20 daily trials in the experimental arrangements For these tests each rat was placed in the apparatus and observed for a period of six minutes The temporal sequences in terms of stimulus presentations and rest intervals were so outlined for the 6-minute period that each time the animal jumped upon or down from the small grid a record could be placed in the time scale Thus, it was possible to determine not only the number of jumps in the absence of sound and shock, but also to note whether these reactions coincided with the time intervals of auditory stimulation The number of would-be correct responses, had the auditory stimulus been presented, was calculated. The trend of these data was so clear that a tabular presentation is considered at this time unnecessary In every instance the reactions of the animal were far below the criterion of learning accepted in these studies. For all animals in terms of both tests the data revealed that chance movements upon or down from the small grid might be expected to coincide with the temporal presentation of the tone in approximately 25 per cent of the time For the buzzer stimulus these responses occurred about 7 per cent of the time The 90 per cent criterion of learning was, therefore, believed to be sufficiently high to preclude the possible explanation of the data in Tables 1 and 2 as due to chance locomotions Further, it is believed that these results would support empirically the arbitrary selection of a lower criterion of learning under the experimental arrangements already described Thus these data indicate that the training given in the experimental arrangements already described has decreased rather than increased the number of diffuse locomotor reactions

The question of the relative difficulty between an ascending and a descending response inherent in the apparatus and training procedure as described may be raised by some In as much as the small grid was

located at a height of three inches above the floor grid it is maintained that any difference in difficulty certainly would not be great Further, jumping up, over, and down reactions are known to be well within the repertoire of activities of both the normal and operated rats. Study of the frequency with which ascending or descending responses were made by normal animals to auditory stimulation during the training periods show wide ranges in variation Survey of the reactions of the median animal indicated an approximately equal number of each of these response types For operated animals the data are in general similar with the exception that a larger percentage of the animals appeared to prefer the descending reaction as shown by their records That these reactions called for in this arrangement are still within the repertoire of activities of operated animals is shown conclusively by Lashley (4) in his series of studies dealing with cerebral mechanisms involved in the rat's solution of latch box problems. The present question of frequency of ascending and descending responses can, of course, be avoided if the investigator wishes In another study, to illustrate this point, four animals were trained in an average of 200 trials to make a sequential locomotor response to a noise stimulus This problem required the acquisition of both the ascending and descending responses in that order. For the present study, however, it is maintained that the question of ascent or descent from the small grid is for general problems of this and allied natures of little significance. It is conceded, however, that for other problems these differential reactions might well be of importance As gross indicators of responses to auditory stimuli it is believed that they are not

In summary of this preliminary and illustrative series of experiments dealing with the problem of training the rat to respond to auditory stimuli, it is concluded that the methodology described above offers a rapid and reliable instrument for the training of small animals to acoustic stimuli It is further suggested at this time that other sense modalities than that of audition might be studied by only slight modifications in the present experimental arrangements

References

1 Brogden, W R, & Culler, E Device for the motor conditioning of small animals *Science*, 1936, 83, 269-270

2. HUNTER, W. A., & PENNINGTON, L. A. A new apparatus and method for training the rat in auditory discrimination problems. *Science*, 1939, 89, 87-88.
3. KRECHEVSKY, I. Brain mechanisms and brightness discrimination learning. *J. Comp. Psychol.*, 1936, 21, 405-455.
4. LASHLEY, K. S. Studies of cerebral function in learning: XI. The behavior of the rat in latch box situations. *Comp. Psychol. Monog.*, 1935, 11, No. 2. Pp. 1-42.
5. MUENZINGER, K. F., & WALZ, F. C. An examination of electrical current-stabilizing devices for psychological experiments. *J. Gen. Psychol.*, 1934, 10, 477-482.
6. PENNINGTON, L. A. The auditory localizing behavior of the white rat in relation to cerebral function. *J. Genet. Psychol.*, 1935, 46, 264-283.
7. ———. The function of the brain in auditory localization. *J. Comp. Psychol.*, 1938, 25, 195-211.

Department of Psychology
University of Illinois
Urbana, Illinois

The Journal of Genetic Psychology, 1940, 57, 461-474.

BOOKS

The *Journal of Genetic Psychology,* the *Journal of General Psychology,* and the *Journal of Social Psychology,* will buy competent reviews at not less than $2 per printed page and not more than $3 per printed page

Conditions Only those books that are listed below in this section are eligible for such reviews In general, any book so listed contains one or more of the following traits. (*a*) Makes an important theoretical contribution, (*b*) consists largely of original experimental research; (*c*) has a creative or revolutionary influence in some special field or the entire field of psychology, (*d*) presents important techniques

The books are listed approximately in order of receipt, and cover a period of not more than three years A reviewer must possess the Ph D. degree or its equal in training and experience

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor Just send in your review It does not matter if the book in question has been reviewed before

(1938)

STRONG, E K Psychological Aspects of Business New York McGraw-Hill, 1938 Pp 629

BOAS, F. The Mind of Primitive Man. (Rev. ed) New York: Macmillan, 1938 Pp 285

SHERMAN, M. Mental Conflicts and Personality. New York: Longmans, 1938 Pp 319

GUTHRIE, E R. The Psychology of Human Conflict New York: Harper, 1938 Pp 408

MURPHY, G., & LIKERT, R Public Opinion and the Individual New York: Harper, 1938 Pp 316

KATZ, D , & SCHANCK, R L Social Psychology New York Wiley, 1938 Pp 700

ELLIS, W D A Source Book of Gestalt Psychology New York Harcourt, Brace, 1938 Pp 403

WOODWORTH, R S Experimental Psychology New York: Holt, 1938 Pp 889.

SEASHORE, C. E Psychology of Music New York McGraw-Hill, 1938 Pp 408.

SKINNER, B F The Behavior of Organisms. New York Appleton-Century, 1938 Pp 457

TERMAN, L M Psychological Factors in Marital Happiness New York McGraw-Hill, 1938 Pp 474

MURRAY, H A , *et al* Explorations in Personality. New York Oxford Univ Press, 1938 Pp 761

LANDIS, C., & PAGE, J. D. Modern Society and Mental Disease. New York: Farrar & Rinehart, 1938. Pp. 190.
MUNN, N. L. Psychological Development. New York: Houghton, Mifflin, 1938. Pp. 582.
RASHEVSKY, N. Mathematical Biophysics. Chicago: Univ. Chicago Press, 1938. Pp. 340.
KOHLER, W. The Place of Value in a World of Facts. New York: Liveright, 1938. Pp. 418.
OGDEN, R. M. The Psychology of Art. New York: Scribner, 1938. Pp. 291.

(1939)

ROSETT, J. The Mechanism of Thought, Imagery, and Hallucination. New York: Columbia Univ. Press, 1939. Pp. 289.
PRATT, C. C. The Logic of Modern Psychology. New York: Macmillan, 1939. Pp. 185.
DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., SEARS, R. R., *et al.* Frustration and Aggression. New Haven: Yale Univ. Press, 1939. Pp. 209.
GESELL, A., CASTNER, B. M., THOMPSON, H., & AMATRUDA, C. S. Biographies of Child Development. New York: Harper, 1939. Pp. 328.
GOLDSTEIN, K. The Organism. New York: World Book, 1939. Pp. 533.
BARTLETT, F. C., GINSBERG, M., LINDGREN, E. J., & THOULESS, R. H. The Study of Society. New York: Macmillan, 1939. Pp. 498.
MAIER, N. R. F. Studies of Abnormal Behavior in the Rat. New York: Harper, 1939. Pp. 81.
WECHSLER, D. The Measurement of Adult Intelligence. Baltimore: Williams & Wilkins, 1939. Pp. 229.
WOODWORTH, R. S. Psychological Issues. New York: Columbia Univ. Press, 1939. Pp. 421.
KARDINER, A. The Individual and His Society. New York: Columbia Univ. Press, 1939. Pp. 503.
STEVENS, B. The Psychology of Physics. Manchester, England: Sherratt & Hughes, 1939. Pp. 282.
ELLIS, H. My Life. Autobiography. New York: Houghton-Mifflin, 1939.

1940

CANTRIL, H. The Invasion from Mars. A Study in the Psychology of Panic. Princeton: Princeton Univ. Press, 1940. Pp. 228.
RASHEVSKY, N. Advances and Applications of Mathematical Biology. Chicago: Univ. Chicago Press, 1940. Pp. 214.
BURLINGAME, L. L. Heredity and Social Problems. New York: McGraw-Hill, 1940. Pp. 369.
DENNIS, W. The Hopi Child. New York: D. Appleton-Century, 1940. Pp. 204.
HILGARD, E. R., & MARQUIS, D. G. Conditioning and Learning. New York: Appleton-Century, 1940. Pp. 429.
HARTMAN, G. W., NEWCOMB, T., *Editors*. Industrial Conflict. A Psychological Interpretation. New York: Cordon, 1940. Pp. 583.
DOOB, L. W. The Plans of Men. New Haven: Yale Univ. Press, 1940. Pp. 411.
HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. B. Mathematico-Deductive Theory of Rôte Learning. New Haven: Yale Univ. Press, 1940. Pp. 329.
LANDIS, C., *et al.* Sex in Development. New York: Hoeber, 1940. Pp. 329.

CRITICAL REVIEWS OF RECENT BOOKS

(Goldstein, K. The Organism. New York: American Book, 1939. Pp. 519.)

REVIEWED BY ROBERT M. OGDEN

This is the author's translation into English of a book first written in German and published in 1934.

In content the book is a comprehensive statement of a holistic, Gestalt biology applied to the behavior of man. While others have shown the need for an organismic interpretation of animal behavior, and likewise for a Gestalt theory of experience, Goldstein bases his conclusions upon a wide psychiatric experience and practice. His interpretations are thus derived both from a knowledge of the functions of the living organism and from the experienced phenomena reported to him by his numerous patients.

The key to Goldstein's interpretation may be found in his use of the term *performance* to describe the total behavior of the organism with overt reference to its environment. Such ordered activity as a performance implies cannot be explained on the basis of reflex behavior. Instead, "the end to which each process normally tends *is determined by its significance for the essential tasks of the whole organism*" (p. 119). Neither an analysis of behavior into reflexes nor a study of the physico-chemical processes of living organisms can explain the performance. The "total excitation pattern *is not confined to a definite anatomical structure, but represents a definite excitation Gestalt* which can utilize for its course any available structure" (p. 232).

The so-called "conditioned reflex" is a phenomenon occasioned by human interference with more natural performances. Though of great interest in training and drill, the natural environment would seldom present such conditions for learning.

> Persons whose right arms are paralyzed or amputated learn very quickly to write with the left. It is not really correct to say "they learn", they do not need to learn because *in principle they know it already at the first trial*. What they really learn is how to overcome obstacles (p. 234). *Localization of a performance no longer means to us an excitation in a certain place, but a dynamic process which occurs in the entire nervous system, even in the whole organism, and which has a definite configuration for each performance* (p. 260).

While behavior is always holistic, it may emphasize performance which is conscious, or the attitude of an inner state of feeling, or the processes of somatic events The effects of former reactions are not remembered when they are no longer a part of our attitudes, but they can be remembered when the individual is brought into a situation similar to the one in which these reactions have belonged

The Freudian notion of sex in childhood with its references to "father," "mother," "child," "ethical norm," "incest," etc., is misleading because sex in childhood is very different from sex in adult life, and the significances of the terms cited are strictly phenomena of adulthood

Not only are the differences marked between the significance of performances in childhood and in adult life; there is a persistent difference between the life of the animal and the life of man. This difference refers to the human "potentiality to focus on the 'possible,' to arrest, so to speak, the world in its course, to picture it, and to shape the coming to terms with the world by virtue of this ability" (p 473)

To one who already accepts an organismic interpretation of life and behavior, this book is a storehouse of supporting material Unfortunately, the organization of the material into a clear-cut exposition of holistic theory is less effective than one might wish Although the criticism of divergent theories, such as those employing mechanistic concepts, is always apt and convincing, one never gets a precise statement of holistic theory References to the systematic contributions of men like Child, Coghill, and Lashley are casual, and there is no reference to the organismic theory developed by Ritter, or to the work of General J. C. Smuts who has made the term *holism* familiar to English readers

The nearest approach to a definite conception of the foundations of behavior is given in terms of flexor and extensor movements. We are told, for instance, that with the help of the flexor performances the human organism

> builds up an understanding of the properties of the stimulus object and eventually incorporates it In the *defense* performances, however, flexor movements hardly participate. At most, they are needed for gaining a superficial knowledge of the dangerous character of the object and the situation The essential reaction type consists of *extensor movements*, through which the object is removed (pp 136-7)

Later we are told that "we find everywhere a *closer relationship of the flexor movements to the cerebral cortex* (and to the cerebellar cortex, which apparently represents only an appendix of the former), and of the extensor movements to the deeper lying apparatuses" (p. 482)

What the author means to say, I take it, is that flexor performances are generally voluntary and conscious, whereas extensor performances are deep seated and have "more reference to the external world" (p 484) than to the self Since, however "not every flexor movement is a flexor performance" (p. 487), one does not know quite what "flexor" and "extensor" really mean in the explanation of human and animal behavior

If instead of translating his book, the author had rewritten it in English, he might have escaped some of the difficulties which arise when an alien mode of thought and exposition is reproduced in another tongue If Goethe's *"Dasein in Tätigkeit"* can be translated as *"Being in Actuality"* (p. 364) one wonders if the author may not have done himself grave injustice throughout by an imperfect rendition of his ideas into English

Department of Psychology
Cornell University
Ithaca, New York

(Goldstein, K. The Organism New York: American Book, 1939 Pp 519)

REVIEWED BY LIVINGSTON WELCH

Very few concepts in psychology or biology are as unlimited in their scope, and uncrystallized as the "organism" Its extreme universality, however, is not necessarily the most perplexing feature. The meaning of such a concept in some instances may be quite adequate, still, its ultimate significance in others may entail enumeration, definition, and interrelationship of much of what the concept implies This is where the difficulty begins The same is true if we start instead at the bottom with any hierarchy of concepts. Identification of one specific, concrete object is a comparatively simple matter Problems do not really arise until one begins to relate and to classify such objects The greatest enigmas in scientific investigation lie in the middle range between these extremes

The study of animal and human organism has been approached from both ends of this scale One method involves an analysis of parts in isolation, while the other consists of the study of the parts in their relation to the natural whole of the organism and its interaction with the environment. Evidence for the dependence of the normal functioning of the part upon the whole is also adduced by studying the changes that take place in the remaining whole when one bodily *part after another is destroyed* or when one functional part after another ceases. The main purpose of this book is to demonstrate the methodological necessity to determine the interrelationship between part and whole. It proposes the argument that often those concepts which we assume to be the simplest turn out to be the most complex For instance, the so-called simplicity that some find in reflex phenomena is at best but the result of isolation in an artificial environment which the investigator sets up or which pathology has produced Under natural conditions the reflex phenomena are much more complex and turn out to be performances in which the entire organism is partaking. Reflex phenomena occur under two sets of conditions, (*a*) under the artificial isolation which either results from pathology of the nervous system or the experimental arrangement by which the remaining whole of the organism is prevented from participating and kept unaltered, (*b*) in border situations where the organism is compelled to adjust suddenly in an

undeliberated automatic fashion. The author gives numerous examples to support this conclusion. It is unlikely that he would deny that any knowledge can be gained from the examination of any part in isolation, but he is in strict agreement with many Gestalt psychologists in maintaining that the interrelationship between parts is as important as the parts themselves, hence, any procedure that would enable the study of both simultaneously would be the most significant and the least misleading.

Dr Goldstein stresses the point that new hypotheses are necessary in order to bring biological facts into accord with one another and raises the question as to whether it is at all feasible "to develop biology on a strictly scientific basis." Immediately following this passage is the statement that his description of the organism is not presented as a theory. His aim, he says, "is not to offer theoretical speculation, but a presentation of the facts themselves, and a discussion of those explanatory concepts which these facts suggest and through which in turn a reliable comprehension of biological phenomena are attainable." Obviously, his use of the term "theory" has a much narrower scope than what we are accustomed to in this country. Hence, any confusion at this point would seem to fall within the realm of translation

This work involves much more than compiling a series of conclusions which are the results of experimentation. It requires adequate categories, clearly defined, which allow for a fruitful interpretation of data. It calls for a general scheme into which the few facts that we know may fit The system Dr. Goldstein has adopted is akin to that of the Gestalt school The distinction of his work lies in its unusual comprehension, together with certain specific differences that carry him beyond the boundaries of the school's orthodoxy. His analysis of the organism extends from the subject of the reflex to the more complex relationships between the organism and its environment.

Like Koffka, he argues against the notion of single reflexes in a chain, and maintains that in many of the characteristics the so-called instincts are more closely related to the whole than the reflexes Instincts, he insists, are processes which belong essentially to the life of the organism "They are not artificial reactions elicited to serve the investigator's purpose, nor are they reactions occasioned by virtue of an inadequate milieu in border situations, both being the case in most reflexes" In the processes of this discussion it

turns out that instinct for Goldstein means a drive, as for example, thirst or hunger, rather than the Gestalt notion of unlearned purposeful behavior

The concept of configuration plays an important part in his interpretation of agonistic and antagonistic innervation, as well as visual stimulation and its resulting response. Concerning the former, he maintains that "the ratio of innervation of agonist and antagonist always depends on the configuration of the whole organism," while in reference to perception he suggests that to a specific color stimulation *corresponds a specific response pattern* of the entire organism.

The disappearance of former reactions in child development, the author explains, is usually described in terms of repression He suggests that the continual formation of new patterns renders ineffective former attitudes. "The factor," he states, "which actuates the so-called repressing is formed neither through prohibition from without, nor by a censor, nor by an ego, nor by a super-ego Rather through maturation, new patterns of the organism are formed conforming to the human species in general and to the cultural pattern of that particular milieu in which the child grows up."

Not only are the Freudian concepts of "repression" and the "ego" criticized, but also the Freudian notion of the "unconscious," which is interpreted as nothing but the entering of a former reaction pattern of the organism into a present response when the situation is suitable The emerging of the so-called unconscious is nothing but the result of strong after effects of certain patterns which have not been sufficiently integrated into the properly centered behavior of the whole organism"

The concepts of "figure" and "ground" play an important part in substituting the distinction between conscious and unconscious mental behavior

Dr. Goldstein lays particular stress on the tendency towards preferred behavior which in many respects corresponds to the tendency toward the good Gestalt. Preferred behavior he considers the essential criterion of ordered behavior and of the constancies of a definite organism. "The operation of this tendency" he says, "includes the so-called pregnanz, the closure phenomena, and many other characteristics of Gestalt. In fact, they are only intelligible from this tendency." Only through the analysis of the interactional totality of the outer and inner fields do the reasons become clear just why a certain pattern, a certain action appears as a "good Gestalt."

He goes on to say that the most reasonable explanation of good Gestalt is not to be in terms of simplicity and minimal energy expenditure, because simplicity is to be defined "on the basis of the demands of the individual task" and the nature of the particular organism confronted with this task. His interpretation of phenomena of pregnanz throws a great deal of light upon this tendency in many situations. The most important contribution seems to be the observation that preferred behavior or good Gestalt can be measured by the criterion of an ordered state in the rest of the organism, whenever preferred behavior occurred in one performance field.

The author devotes an entire chapter to an explanation of the points in which he differs from the Gestalt school He maintains that his guiding principle is somewhat different from that of the Gestalt psychologists inasmuch as the whole, the Gestalt, has always meant to him the whole organism, and not the phenomena in one field or merely introspective experience.

This would indicate that his use of the concept "pregnanz" must not be restricted to the narrow limits in which it is found isolated at a moment in one portion of an organism's visual field. The Goldstein "pregnanz" by definition implies the behavior of the entire organism at the moment. The point is made still clearer in the quotation—"preferred behavior, good Gestalt, or whatever one chooses to call it, represents a very definite form of coming to terms of the organism with the world, that form in which the organism actualizes itself according to its nature in the best way" Of course, the phrase "best way" is open to further inquiry, and raises the question whether it is scientifically possible to develop objective criteria of what is "adequate" for an organism. In answer Dr Goldstein evolves the criterion of "ordered behavior" as contrasted with "disordered" or "catastrophic" behavior, the experiential complement of which is anxiety. An organism is in an ordered state and performs in the best possible way, or realizes its nature at its best under conditions of preferred behavior, i e., when a performance fulfills the criteria of good Gestalt and if, at the same time, the remaining whole of the organism is in ordered condition and not deranged Hence, a performance which may appear to conform with the characteristic of a good Gestalt but is attained at the expense of order in the remainder of the organism does not present the best way of actualizing the organism's nature. Such a constellation could only be justified in the fulfillment of the present task, in

spite of discomfort it is meaningful for the organism's constitution as a whole at a distant time.

A lengthy discussion of his main principle is followed by minor treatments of his concept of "physical Gestalt," individual field forces, and the functional significance of the latter On such points he differs from the ordinary Gestalt interpretations, but these, he says, do not offer insurmountable discrepancies between his own theory and that of the Gestalt school.

This book is unquestionably a scholarly and extremely stimulating treatment of the broadest problem one can attempt to study in the field of biology or psychology—a work that in many respects is unparalleled in the field of contemporary psychological interpretation. The present volume is an English translation of the German edition published in 1934

Department of Psychology
Hunter College
2 Park Avenue
New York City

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i e, the most recently received books will be found at the end of the list)

ROETHLISBERGER, F J, & DICKSON, W. J. Management and the Worker. Cambridge: Harvard Univ Press, 1939 Pp 615

JANNEY, J E Laboratory Workbook in Psychology New York: Harper, 1939 Pp. 154

RATNER, J (Editor) Intelligence in the Modern World: John Dewey's Philosophy New York Random House, 1939. Pp. 1075.

GARRISON, K C The Psychology of Exceptional Children. New York: Ronald Press, 1940 Pp. 351.

RASHEVSKY, N Advances and Applications of Mathematical Biology Chicago: Univ. Chicago Press, 1940. Pp 214.

FRAMPTON, M E, & ROWELL, H G. (Editors). Education of the Handicapped Vol. 2 Problems. New York: World Book, 1940 Pp 440.

SELLING, L S Men Against Madness. New York Greenberg, 1940. Pp. 342.

ERDELYI, M., & GROSSMAN, F (Editors). Dictionary of Terms and Expressions of Industrial Psychology New York: Pitman, 1939. Pp 98.

HOHMAN, L. B As the Twig is Bent New York Macmillan, 1940 Pp 291

KUNKEL, F, & DICKERSON, R. E How Character Develops New York: Scribner, 1940 Pp. 274

GRABBE, P. We Call It Human Nature. New York Harper, 1939 Pp. 120.

VALENTINE, W L. Student's Guide for Beginning the Study of Psychology. (Revised edition.) New York Prentice-Hall, 1940 Pp 249

GUILFORD, J. P (Editor) Fields of Psychology. New York Van Nostrand, 1940 Pp 695.

TEAGARDEN, FLORENCE M Child Psychology for Professional Workers. New York: Prentice-Hall, 1940. Pp 641

DURBIN, E. F. M., & BOWLBY, J Personal Aggressiveness and War New York: Columbia Univ Press, 1939 Pp 154

BURLINGAME, L. L. Heredity and Social Problems. New York: McGraw-Hill, 1940 Pp 369

WHEELER, R H The Science of Psychology. New York: Crowell, 1940 Pp 436.

CANTRIL, H The Invasion from Mars A Study in the Psychology of Panic Princeton: Princeton Univ Press, 1940 Pp 228.

LATTIMORE, O Inner Asian Frontiers of China New York, Amer Geograph Soc, 1940 Pp. 585

GUILFORD, J. P, & WALTON, W. E Studies in Elementary Psychology. New York Van Nostrand, 1940 Pp 266

BAKER, H J The Art of Understanding Boston: Christopher, 1940 Pp 400.

MERRY, F K., & MERRY, R V. From Infancy to Adolescence New York, 1940 Pp. 330

RHINE, J B, ET AL. Extra Sensory Perception after Sixty Years. New York Holt, 1940. Pp. 463

WARDEN, C J., JENKINS, T N, & WARNER, L H Comparative Psychology: II Plants and Invertebrates New York: Ronald Press, 1940 Pp. 1070

WANG, C K A An Annotated Bibliography of Mental Tests and Scales. Peiping. Catholic Univ Press, 1939 Pp 725

BERNHEIMER, R., CARPENTER, R., KOFFKA, K., & NAHM, M. C. Art: A Bryn Mawr Symposium. Bryn Mawr: Bryn Mawr College, 1940. Pp. 350.
DOOB, L. W. The Plans of Men. New Haven: Yale Univ. Press, 1940. Pp. 411.
BODE, B. H. How We Learn. Boston: Heath, 1940. Pp. 308.
HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., & FITCH, F. B. Mathematico-Deductive Theory of Rote Learning. New Haven: Yale Univ. Press, 1940. Pp. 329.
HARTMAN, G. W., & NEWCOMB, T., *Editors*. Industrial Conflict: A Psychological Interpretation. New York: Cordon, 1940. Pp. 583.
KENDIG, I., & RICHMOND, W. V. Psychological Studies in Dementia Praecox. Ann Arbor, Mich.: Edwards, 1940. Pp. 211.
SIMMONS, R. McK. A Study of a Group of Children of Exceptionally High Intelligence Quotient in Situations Partaking of the Nature of Suggestion. New York: Teach. Coll., Columbia Univ., 1940. Pp. 112.
KATONA, G. Organizing and Memorizing. New York: Columbia Univ. Press, 1940. Pp. 318.
HUSBAND, R. W. General Psychology. New York: Farrar & Rinehart, 1940. Pp. 513.
HILGARD, E. R., & MARQUIS, D. G. Conditioning and Learning. New York: Appleton-Century, 1940. Pp. 429.
GRIFFITH, E. S. The Impasse of Democracy. New York: Harrison-Hilton, 1940. Pp. 380.
TIFFIN, J., KNIGHT, F. B., & JOSEY, C. C. The Psychology of Normal People. Boston: Heath, 1940. Pp. 512.
SMITH, G. M. Workbook in Psychology. New York: Holt, 1940. Pp. 134.
ANDERSON, J. P. A Study of the Relationships between Certain Aspects of Parental Behavior and Attitudes and the Behavior of Junior High School Pupils. New York: Teach. Coll., Columbia Univ., 1940. Pp. 196.
BENZ, MARGARET G. Family Counseling Service in a University Community. New York: Teach. Coll., Columbia Univ., 1940. Pp. 125.
GENTRY, J. R. Immediate Effects of Interpolated Rest Periods on Learning Performance. New York: Teach. Coll., Columbia Univ., 1940. Pp. 57.
GANS, R. Critical Reading Comprehension in the Intermediate Grades. New York: Teach. Coll., Columbia Univ., 1940. Pp. 135.
OSBORN, F. Preface to Eugenics. New York: Harper, 1940. Pp. 312.
WERNER, H. Comparative Psychology of Mental Development. New York: Harper, 1940. Pp. 510.
SORENSON, H. Psychology in Education. New York: McGraw-Hill, 1940. Pp. 489.
LAFORGE, R. The Relativity of Reality. Nerv. & Ment. Dis. Monog., 1940, No. 66. Pp. 92.
THOMSON, G. H. An Analysis of Performance Test Scores of a Representative Group of Scottish Children. London: Univ. London Press, 1940. Pp. 58.
SEELEY, H. F., & HACKETT, W. A. Experiences in Speaking. Chicago: Scott, Foresman, 1940. Pp. 512.
BENEDICT, RUTH. Race: Science and Politics. New York: Modern Age, 1940. Pp. 274.
LANDISS, C., *et al.* Sex in Development. New York: Hoeber, 1940. Pp. 329.
EATON, HELEN S. Semantic Frequency List. Chicago: Univ. Chicago Press, 1940. Pp. 441.

VOLUME 2[illegible]—[illegible]

[illegible]

VOLUME 21—February and May, 194[illegible]

[illegible]

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts